The Continuum Random Tree II: An Overview

David Aldous[*]

University of California, Berkeley

1 INTRODUCTION

Many different models of random trees have arisen in a variety of applied setting, and there is a large but scattered literature on exact and asymptotic results for particular models. For several years I have been interested in what kinds of "general theory" (as opposed to *ad hoc* analysis of particular models) might be useful in studying asymptotics of random trees. In this paper, aimed at theoretical probabilists, I discuss aspects of this incipient general theory which are most closely related to topics of current interest in theoretical stochastic processes. No prior knowledge of this subject is assumed: the paper is intended as an introduction and survey.

To give the really big picture in a paragraph, consider a tree on n vertices. View the vertices as points in abstract (rather than d-dimensional) space, but let the edges have length ($= 1$, as a default) so that there is metric structure: the distance between two vertices is the length of the path between them. Consider the average distance between pairs of vertices. As $n \to \infty$ this average distance could stay bounded or could grow as order n, but almost all natural random trees fall into one of two categories. In the first (and larger) category, the average distance grows as order $\log n$. This category includes supercritical branching processes, and most "Markovian growth" models such as those occurring in the analysis of algorithms. This paper is concerned with the second category, in which the average distance grows as order $n^{1/2}$. This occurs with Galton-Watson branching processes conditioned on total population size $= n$ (in brief, CBP(n)). At first sight that seems an unnatural model, but it turns out to coincide (see section 2.1) with various combinatorial models, and is similar to more general models of *critical* branching processes conditioned to be large (in any reasonable way). The fundamental

---

[*] Research supported by N.S.F. Grants MCS87-01426 and MCS 90-01710

fact is that, by scaling edges to have length $n^{-1/2}$, these random trees converge in distribution as $n \to \infty$ to a limit we call the CCRT (for *compact continuum random tree*). This was treated explicitly in Aldous [2] in a special case and in Aldous [3] in the natural general case, though (as we shall see) many related results are implicit in recent literature. Thus asymptotic distributions for these models of discrete random trees can be obtained immediately from distributions associated with the limit tree. The limit tree is closely connected with *Brownian excursion.* In fact two different 1-parameter processes associated with the tree – the *search depth* process and the *height profile* process – are intimately connected with Brownian excursion (sections 2.4 and 3.2). Section 2 is a chatty account of 4 different ways of looking at the CCRT. In section 3 I take natural distributional questions about CBP(n) asymptotics (with known or unknown answers), which can be expressed in terms of the CCRT and see what can be said about the limit distributions, using the Brownian excursion representation in particular. Nothing I say is essentially new: I use the word "novel" (intended to be weaker than "new") to refer to results about CBP(n) asymptotics obtainable from known Brownian excursion results (e.g. Corollaries 3 and 6, and Proposition 12) and vice versa (e.g. (41) as a fact about Brownian excursion). One could conversely pick haphazardly some facts about Brownian excursion and apply them to random trees, but that somehow seems less interesting.

Scaling the edges of CBP(n) to have length $n^{-\alpha}$ $(0 < \alpha < 1/2)$ gives (section 2.5) another limit tree I call the SSCRT (*self-similar continuum random tree*). Further, the same limit tree is obtained whether we root at the progenitor or whether we re-root at a uniform random individual in the population. This limit tree – which relates to the 3-dimensional Bessel process BES(3) in the same way that the CCRT relates to Brownian excursion – is less natural from the combinatorial viewpoint. But being more tractable (from the self-similarity inherited from BES(3)) it is useful in the theoretical stochastic process investigations below.

Sections 5 and 6 are speculative. There has been recent theoretical interest in existence, uniqueness and properties of "Brownian motion" whose state space is some deterministic fractal set in d dimensions, the set typically constructed by some recursive procedure giving strong regularity properties. Our limit trees are "dimension 2" (inherited from Brownian sample paths), and it is intuitively clear that "Brownian motion" can be defined with these trees as its state space. Unlike other exotic state spaces, we can actually do some simple distributional calculations with these Brownian motions, and the purpose of section 5 is to present these back-of-an-envelope calculations. To develop rigorously a theory of Brownian motion on general continuum trees would be an interesting project, and some thoughts are presented in section 5.2.

Section 6 is a quixotic venture into superprocesses. It is trivial to construct Markov processes *indexed by* a continuum tree. Making the index set the particular CCRT or SSCRT gives variants of the usual superprocess. This is the idea developed by theoreticians under the name "historical process", but the theoretical literature makes this appear a deep and sophisticated object. I assert one should start from scratch and regard a superprocess as a tree-indexed process rather than as a measure-valued process. My purpose is to indicate (section 6.1) how this leads to insights which seem simpler or different from those obtained in the traditional approach.

Acknowledgements. I thank Jim Pitman for much help with and information about Brownian excursion and BES(3), Steve Evans for help with superprocesses, and Martin Barlow for discussions on diffusions on fractals.

2 THE BIG PICTURE

The first four subsections elaborate on the following four fundamental facts.

- Conditioned Galton-Watson branching processes correspond to a natural and well-studied class of combinatorial models of random trees.
- One particular model can be constructed from simple random walk conditioned on first return to 0 at time $2n$, and so its asymptotics can be expressed in terms of Brownian excursion.
- Another particular model can be constructed from a direct (i.e. not involving conditioning) algorithm, and by taking limits one gets a direct algorithm for global construction of a limit tree.
- By considering asymptotics of subtrees spanned by a fixed number of randomly chosen vertices, one sees that the limit random tree must be the same (up to a scale factor) for all models in the class.

Foundational work giving rigorous definitions and proofs concerning existence of "continuum trees" (without any specific probability model present) and abstract convergence results is in Aldous [3], and it is not worth repeating such "general abstract nonsense" here.

2.1 CBP(n) and Combinatorial Models

Let $\xi \geq 0$ be integer-valued and satisfy

$$E\xi = 1$$

$$0 < \mathrm{var}\ \xi = \sigma^2 < \infty. \tag{1}$$

Such a ξ is d-lattice, for some $d \geq 1$. We want to allow $d > 1$ for natural combinatorial examples (e.g. binary trees). Associate with ξ the distribution $\hat{\xi}$ defined by

$$P(\hat{\xi} = i) = (i+1)P(\xi = i+1),\ i \geq 0 \tag{2}$$

and note that

$$E\hat{\xi} = E\xi(\xi - 1) = \mathrm{var}\ \xi = \sigma^2.$$

Consider the simple Galton-Watson branching process with offspring distribution ξ, starting with 1 individual in generation 0. Write $\mathcal{T}$ for the "family tree" of this branching process. Let $\mathcal{T}_n$ have the distribution of $\mathcal{T}$ conditioned on the total population size $|\mathcal{T}| = n$. This CBP(n) (for "conditioned branching process") distribution is our object of study.

Tangential remarks. 1. If ξ and η come from the same exponential family, i.e. for some (c, θ)

$$P(\xi = i) = c\theta^i P(\eta = i);\ i \geq 0$$

then the conditioned branching processes constructed from ξ and from η are identical. Thus we lose no generality by considering only *critical* branching processes. The chance that the total population size is exactly n decreases exponentially fast for sub- and super-critical branching processes, but only polynomially fast in the critical case: in this sense the critical case is most natural as a model for n-trees.

2. In the language of freshman statistics, if ξ is "number of daughters of a randomly-picked mother", then $\hat{\xi}$ at (2) is "number of sisters of a randomly-picked girl". The two distributions are identical iff they are the Poisson(1) distribution.

3. I use "Galton-Watson process" to mean the family tree of the process. Old-fashioned textbooks use it to mean the process of population sizes in successive generations, which I call the "height profile" of the Galton-Watson process.

Simply generated trees. Results about CBP(n) appear in the combinatorial literature under this name (introduced by Meir and Moon [35], apparently unaware of the branching process connection). Though the identification has subsequently become well known, there seems no convenient "translation guide" in existence, so I give one here.

A "rooted" tree simply has one vertex distinguished and called the root: imagine a family tree of descendants of a single progenitor, the root. We consider only rooted trees. Such a tree is called *ordered* if we distinguish birth order: if an individual (vertex) has 3 offspring then these are distinguished as "first", "second" and "third". Consider the family tree $\mathcal{T}$ of the unconditioned Galton-Watson branching process with offspring ξ. Write $p_i = P(\xi = i)$. Then the distribution of $\mathcal{T}$ on rooted ordered trees t is

$$\begin{aligned} P(\mathcal{T} = t) &= \prod_{v \in t} p_{d(v,t)} \\ &= \prod_{i \geq 0} p_i^{D_i(t)} \\ &= \omega(t) \text{ say} \end{aligned} \tag{3}$$

where $d(v,t)$ is the out-degree (number of children) of vertex v in t, and $D_i(t)$ is the number of vertices in t with out-degree i. Thus the distribution of the CBP(n) tree $\mathcal{T}_n$ is specified by

$$P(\mathcal{T}_n = t) \text{ is proportional to } \omega(t) \text{ on } \{t : |t| = n\} \tag{4}$$

where $|t|$ denotes the number of vertices in t.

One can get to (4) without explicitly mentioning Galton-Watson processes. Let $(c_i; i \geq 0)$ be non-negative constants with $c_0 = 1$, and let

$$\phi(y) = \sum_i c_i y^i$$

be the associated generating function. Let $\hat{\omega}(t)$ be some collection of non-negative "weights" for trees. Define

$$y_n = \sum_{t:|t|=n} \hat{\omega}(t)$$

and let $Y(x) = \sum_n y_n x^n$ be the associated generating function. Then it is easy to see the following are equivalent.

$$Y(x) \equiv x\phi(Y(x)) \tag{5}$$

$$\hat{\omega}(t) = \prod_{i \geq 0} c_i^{D_i(t)} \tag{6}$$

A combinatorial definition of "simply generated tree" is "a family of weights satisfying (5), or equivalently (6)". So a random simply generated tree $\mathcal{T}_n$ is defined as

$$P(\mathcal{T}_n = t) \text{ is proportional to } \hat{\omega}(t) \text{ on } \{t : |t| = n\}.$$

To see why this is really the same as the CBP(n) model, note that for any τ with $\phi(\tau) < \infty$ we can define a probability distribution

$$P(\xi = i) = p_i = c_i\tau^i/\phi(\tau), \ i \geq 0. \tag{7}$$

Choose the τ which makes $E\xi = 1$. For $\omega(t)$ defined at (3), we see

$$\omega(t) = \hat{\omega}(t)\ \tau^{|t|-1}/\phi^{|t|}(\tau).$$

Thus on $\{t : |t| = n\}$, ω is proportional to $\hat{\omega}$, and so the two models for $\mathcal{T}_n$ are identical.

Elementary calculations from (7) show that the condition "$E\xi = 1$" specifying τ is the condition

$$\tau\phi'(\tau) = \phi(\tau)$$

and that the variance $\sigma^2 \equiv \text{var}(\xi)$ is

$$\sigma^2 = \tau^2\phi''(\tau)/\phi(\tau). \tag{8}$$

The right-side expression appears in combinatorial papers without mention of its simple interpretation as "offspring variance".

Examples. The idea of all the combinatorial examples is that all n-vertex trees of a certain type should be equally likely. One aspect of "type" is that we can place restrictions on out-degrees. Another aspect is that sometimes we want to distinguish birth-order (*ordered trees*) and sometimes we don't. In the set-up above, ordered trees become the case

$$c_i = 1 \text{ if } i \text{ is an allowed out-degree}, \quad = 0 \text{ if not}$$

and unordered trees become the case

$$c_i = 1/i! \text{ if } i \text{ is an allowed out-degree}, \quad = 0 \text{ if not} .$$

Various offspring distributions $p_i = P(\xi = i)$ are recorded below as a handy reference: the values of σ are needed to connect our results with those in the combinatorial literature on special models. To reiterate the point: the uniform distribution on the following "types of n-vertex tree" coincides with the CBP(n) description with the stated offspring distribution.

ordered (= planar) trees.

Unrestricted degree: shifted geometric distribution $p_i = 2^{-i}, i \geq 0$; $\sigma^2 = 1$.
Strict binary (0 or 2 offspring): $p_0 = p_2 = 1/2$; $\sigma^2 = 1$.

Strict t-ary (0 or t offspring): $p_0 = 1 - 1/t, p_t = 1/t$; $\sigma^2 = t - 1$.
Unary-binary (0, 1 or 2 offspring): $p_0 = p_1 = p_2 = 1/3$; $\sigma^2 = 2/3$.

unordered labelled trees.
Unrestricted degree: Poisson distribution $p_i = e^1/i!, i \geq 0$; $\sigma^2 = 1$.
Unary-binary: $p_0 = \frac{1}{2+\sqrt{2}}, p_1 = \frac{\sqrt{2}}{2+\sqrt{2}}, p_3 = \frac{1}{2+\sqrt{2}}$; $\sigma^2 = \frac{2}{2+\sqrt{2}}$.
Strict t-ary: same as ordered case.

Remark. I have slid over one issue: in the combinatorial story the trees are regarded as rooted and *labelled*, i.e. the n vertices are distinguishable. The distinction between labelled and unlabelled is irrelevant for ordered rooted trees (because the ordering serves to distinguish vertices anyway) but relevant for unordered trees. The model "all unordered unlabelled trees equally likely" does *not* fit into this set-up, and no simple probabilistic description is known.

2.2 Ordered Trees and Brownian Excursion

With a finite rooted ordered tree t on n vertices we can associate the following two sequences (the terminology is not standard).

The *height profile* $(h(j); j \geq 0)$, where $h(j)$ is the number of vertices at distance j from the root.

The *search depth* $(x(i); 1 \leq i \leq 2n - 1)$ defined as follows. At each vertex v, suppose the edges at v leading away from the root are *ordered* as "first", "second", etc. Then depth-first search of the tree is the following deterministic walk $(v(i) : 1 \leq i \leq 2n - 1)$ around the vertices. Let $v(1) =$ root. Given $v(i)$ choose (if possible) the first (in the ordering) edge at $v(i)$ leading away from the root which has not already been traversed, and let $(v(i), v(i+1))$ be that edge. If not possible, let $(v(i), v(i+1))$ be the edge from $v(i)$ leading towards the root.

This walk terminates with $v(2n - 1) =$ root, having traversed each edge exactly once in each direction. Finally, define the search depth $x(i) =$ distance from root to $v(i)$.

There is a connection between the two sequences: for $j \geq 1$

$$h(j) = \text{ number of upcrossings of } (j-1, j) \text{ by the sequence } x(i). \qquad (9)$$

For a random tree distributed as CBP(n) these become random sequences

$(H(j))$ and $(X(i))$, say. Define the *rescaled cumulative height profile process*

$$H^n_t = n^{-1} \sum_{j \le n^{1/2}t} H(j),\ t \ge 0 \tag{10}$$

and the *rescaled search depth process*

$$X^n_t = n^{-1/2} X([2nt]),\ 0 \le t < 1. \tag{11}$$

Conventions about rescaling constants are awkward – e.g. one might want to rescale by $(2n)^{-1/2}$ in (11) – but my conventions are chosen to make rescaled edge-lengths $= n^{-1/2}$ consistently.

Returning to the unscaled process $X(i)$, set $X(0) = X(2n) = -1$. For any model, the process $X(i)$ has steps ± 1 and first returns to the starting level after step $2n$. The simplest model for such a random process would be "simple symmetric random walk, conditioned on first return to starting level at time $2n$". The key fact is that this describes the depth search process in one particular model of random trees: the combinatorial model of "uniform ordered trees", which is the CBP(n) model with shifted geometric (1/2) offspring distribution.

Various forms of this fact have been known to combinatorialists for a long time But its significance for probabilistic asymptotics was overlooked until recently (I learned it from Durrett et al [17], who attribute it to Harris). It is intuitively obvious (and true [16] – see also [13] and [9] p. 104 for references and history) that conditioned random walk rescales to Brownian excursion, and so (for this special model of random trees) the rescaled search depth process converges to Brownian excursion. It is equally intuitively obvious from (9) that the rescaled height profile process converges to the total occupation density of Brownian excursion.

On a finite tree, the search depth process determines the ordered tree in a simple way: each $+1$ step draws a new edge, and each -1 step retraces an existing edge toward the root. So it is intuitively clear that, for the special "uniform ordered n-tree" model, there is a limit tree whose realizations can be constructed from realizations of sample paths $f(t)$ of Brownian excursion. In non-standard terms, an infinitesimal positive increment of f draws an infinitesimal new edge, and an infinitesimal negative increment of f retraces an existing edge toward the root. In standard terms, given $0 < t_1 < t_2 < \ldots < t_k < 1$, let $s_i = \min_{t_i < t < t_{i+1}} f(t)$. Draw an edge of length $f(t_1)$, and label one end "root" and the other end "t_1". Inductively, from t_i move back distance $f(t_i)-s_i$ toward the root, then make a new edge of length $f(t_{i+1})-s_i$ and label its endpoint "t_{i+1}". The shapes of these trees are consistent as (t_i) varies, and define a "continuum tree" with vertices labelled by $0 < t < 1$.

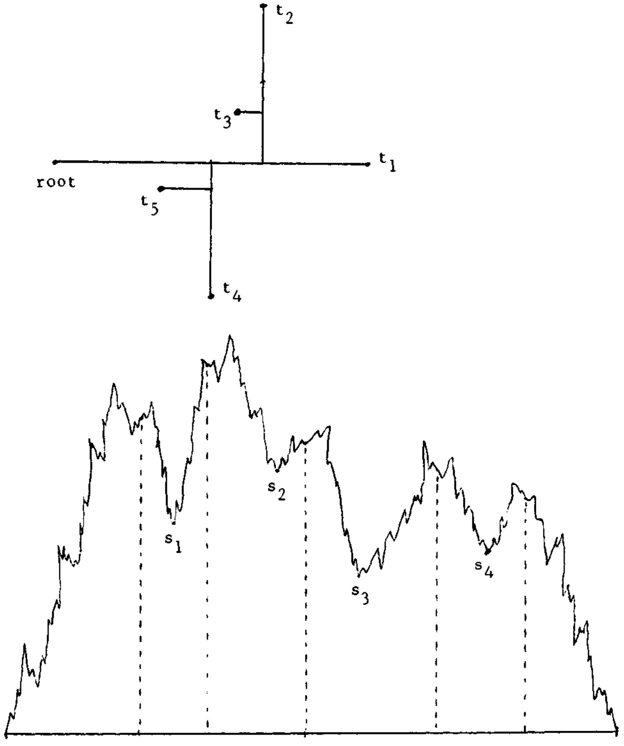

figure 1

A rigorous treatment of constructing continuum trees from continuous functions is given in [3], Theorem 13: it turns out that distributional properties of Brownian excursion are irrelevant, and that any continuous function f_0 with certain qualitative properties (e.g. local minima are dense) can be used. Later we shall use $(\mathcal{S}_0, \mu_0)$ to denote the continuum tree constructed from such a f_0. Regard $\mathcal{S}_0$ as the vertex-set, labelled by $0 < t < 1$, and regard μ_0 as the "uniform probability distribution" on $\mathcal{S}_0$ induced from Lebesgue measure. Write $(\mathcal{S}, \mu)$ for the particular continuum random tree ("the CCRT") constructed from $(2\times)$ Brownian excursion.

2.3 The Limit Trees: Global Constructions

Another special case of CBP(n) is where the offspring distribution is Poisson(1). Combinatorially, this is the uniform random unordered labelled tree. Many algorithms for simulating this random tree are known: the following was discovered in Aldous [5].

Algorithm 1 *Fix* $n \geq 2$.
Take a root vertex 1.
For $2 \leq i \leq n$ *connect vertex* i *to vertex* $V_i = \min(U_i, i-1)$, *where* $U_2, \ldots, U_n$ *are independent and uniform on* $1, \ldots, n$.
Randomly permute the labels.

The advantage of this particular algorithm is that the $n \to \infty$ limit behavior is intuitively easy to see. It is proved in Aldous [2] that the first process (the CCRT) described informally below is the limit when edges are rescaled to length $n^{-1/2}$, and the second process (the SSCRT) is the limit when edges are rescaled to length $n^{-\alpha}$, $0 < \alpha < 1/2$, or more generally to length $1/a(n)$ where $a(n) \to \infty, a(n) = o(n^{1/2})$.

The compact continuum random tree $(\mathcal{S}, \mu)$.
Take a half-line $[0, \infty)$, and cut-and-paste as follows. Let $C_1, C_2, \ldots$ be the times of a non-homogeneous Poisson process of rate $r(t) = t$. Cut the half-line into intervals $[C_i, C_{i+1})$. Start with the line segment $[0, C_1)$, and make 0 the root. Grow a tree inductively by adding $[C_i, C_{i+1})$ as a branch connected to a random point J_i, chosen uniformly over the existing tree. The process is the closure of the union of all branches. □

The self-similar continuum random tree $(\mathcal{R}, \nu)$.
Start at time 0 with an infinite continuous line $[0, \infty)$, and make 0 the root. At time $0 < t < \infty$ there is a tree composed of the original line and of finite line segments connected with each other; only a finite number of such segments connecting with each finite interval of the original line. The process grows according to the rules
(i) in each time increment $(t, t+dt)$, in each segment $(x, x+dx)$ of the tree constructed at time t, there is chance $dt\,dx$ of a "birth";
(ii) if a birth occurs at time t and place x, then a new branch with random exponential(rate t) length is instantly attached at x.
The process is the closure of the tree at time infinity. □

In these limit processes, regard $\mathcal{S}$ and $\mathcal{R}$ as random sets, indicating the spatial position of the limit "continuum tree". Then μ and ν are random measures supported by $\mathcal{S}$ and $\mathcal{R}$, representing how the vertices are spread over the tree.

In other words, with the tree $\mathcal{T}_n$ constructed by Algorithm 1 we associate the empirical distribution μ_n of the vertices: μ_n puts mass $1/n$ on each vertex. As space-rescaled $\mathcal{T}_n$ converges to $\mathcal{S}$, so does μ_n (with the induced space-rescaling) converge to μ. Similarly, when edge-lengths of $\mathcal{T}_n$ are rescaled to length $1/a(n)$ to get the limit $\mathcal{R}$, let ν_n be the measure putting mass $1/a^2(n)$ on each vertex: then ν_n, with the induced space-rescaling, converges to ν.

As the notation suggests, we shall see below that the CCRT $(\mathcal{S}, \mu)$ constructed above is the same as that constructed from (2×) Brownian excursion in the previous section.

2.4 The Convergence Result for CBP(n)

The results in the previous two sections depended on exact combinatorial relations for finite n, in the two special cases. A natural first step in seeking to generalize is to consider the general CBP(n) model. Neither of the previous methods works: there is in general no constructive algorithm like Algorithm 1 known, and while any tree can be coded as a walk with steps ±1 as in section 2.2, the process obtained from general CBP(n) does not have any standard dependence structure which makes convergence to Brownian excursion look easy to prove. But using different techniques (outlined below), an abstract result "rescaled general CBP(n) converges to the CCRT" is proved in Aldous [3] Theorem 23. Without setting up the precise statement of the abstract result, let us state the concrete consequence (which actually turns out to be equivalent to the *a priori* stronger abstract result – c.f. [3]) for the rescaled search depth process X^n at (11).

Theorem 2 *For CBP(n), as $n \to \infty$*

$$(X^n_t; 0 \leq t \leq 1) \xrightarrow{d} (2\sigma^{-1} W_t; 0 \leq t \leq 1)$$

where $W = (W_t; 0 \leq t \leq 1)$ is standard Brownian excursion.

Here "convergence in distribution" is the usual weak convergence of processes. Note we use "Brownian excursion" to mean Brownian excursion *of duration* 1.

An immediate corollary is the result for the rescaled cumulative height profile process H^n at (10).

Corollary 3 *For CBP(n), as $n \to \infty$*

$$(H^n_s; s \geq 0) \xrightarrow{d} (H_{\sigma s/2}; s \geq 0) \tag{12}$$

where

$$H_s = \int_0^1 1_{(W_t \le s)} dt.$$

To use an old-fashioned term, the abstract result behind Theorem 2 is an *invariance principle*: the distribution of the limit tree $\mathcal{S}$ doesn't depend on the offspring distribution ξ, except through the s.d. σ as a scale factor. (This may be thought surprising – one's first guess might be that σ would affect the *shape* of the tree). As with the classical invariance principle (convergence of i.i.d. partial sums to Brownian motion) one might expect the result to be true for much more general models, and we discuss this briefly in section 4.

A final ingredient of the big picture is an "intrinsically tree-ish" description of the CCRT $\mathcal{S}$. To give this, we need to introduce a different species of tree t. Let t have k labelled leaves, a root with degree 1, and binary branchpoints (and hence $2k-1$ edges). Let the edge-lengths be positive reals, and regard the tree as unordered. Such a tree t can be specified by its topological shape t^*, say, and by the $2k-1$ edge-lengths (l_i). Define $\mathcal{R}(k)$ to be a random tree of this type with density

$$f(t^*, l_1, \dots, l_{2k-1}) = s\exp(-s^2/2), \ s = \sum_{i=1}^{2k-1} l_i. \tag{13}$$

In other words, the edge-lengths are independent of the shape of the tree, which is uniform on all shapes; moreover the edge-lengths are exchangeable (and hence we didn't need to specify exactly which edge was edge i). These random trees satisfy the natural consistency condition in k. It turns out that the distribution of $(\mathcal{S}, \mu)$ is specified by the fact that the subtree $\mathcal{R}(k)$ spanned by k "uniform" (i.e. chosen according to μ) random vertices has density (13). More generally, just as ordinary stochastic processes can be specified via consistent families of f.d.d.'s, so ([3] Theorem 3) a random continuum tree can be specified by "random f.d.d.'s", the subtrees spanned by randomly-chosen vertices. The point is that in general there is no "canonical" way of labelling vertices of continuum trees, so random f.d.d.'s are a natural substitute for ordinary f.d.d.'s.

The proof in [3] that rescaled CBP(n) converges is based upon convergence of random f.d.d.'s. Fix k, choose at random k vertices from CBP(n), consider the subtree spanned by these k vertices and the root, rescale and let $n \to \infty$. Using classical asymptotics for sizes of critical BPs it can be shown that the limit tree is (a scale factor σ^{-1} times) $\mathcal{R}(k)$. In [3] we develop such "exchangeability and weak convergence" techniques as a hopefully useful way of establishing convergence of more general models to more general limit continuum random trees. In principle one could seek to prove Theorem 2 directly, by first proving convergence of finite dimensional distributions $(X^n_{t_1}, \ldots, X^n_{t_k})$. In a special model of binary trees this was recently carried out by Gutjahr and Pflug [26], based on exact combinatorial formulas, but the general CBP(n) model seems less tractable. Direct approaches to Corollary 3 are easier, but not powerful enough to establish the Theorem.

We now tie this up with the special constructions of $\mathcal{S}$ in sections 2.2 and 2.3. The construction in section 2.2 of a tree from a function f and points (t_i), applied to a sample path of $2W$ and to k uniform random points, plainly must give a tree isometric to $\mathcal{R}(k)$. The connection with the global construction in section 2.3 is more surprising. Let $\hat{\mathcal{R}}(k)$ be the subtree obtained from the first k branches $[C_{i-1}, C_i], i \leq k$ in the global construction. Then a direct computation ([3] section 4.3) shows that $\hat{\mathcal{R}}(k)$ has distribution (13):

$$\hat{\mathcal{R}}(k) \overset{d}{=} \mathcal{R}(k) \tag{14}$$

In other words, $\hat{\mathcal{R}}(k)$ is isometric to the subtree $\mathcal{R}(k)$ of $\mathcal{S}$ spanned by k *randomly* chosen vertices of $\mathcal{S}$.

It is intuitively clear that the natural "local" result associated with Corollary 3 should be true. Write $H(j)$ for the number of vertices at height j from the root, and define the rescaled height profile process

$$h^n_s = n^{-1/2} H([n^{1/2}s]).$$

Conjecture 4 *For CBP(n), as $n \to \infty$,*

$$(h^n_s; s \geq 0) \xrightarrow{d} (\frac{\sigma}{2} l_{\sigma s/2}; s \geq 0)$$

where

$$l_s = \frac{d}{ds} \int_0^1 1_{(W_t \leq s)} dt$$

is the total occupation density of Brownian excursion W.

The "weak convergence" methods of [3] are too weak to be used here. In some special cases there are exact combinatorial expressions for means and moments of $H(j)$, and in these special cases one could no doubt establish Conjecture 4, but the general case seems to require delicate analytical asymptotics. Distributional properties are discussed in section 3.2.

Remark: local time convention. Above I use total occupation density for Brownian excursion, and later I use total occupation density for BES(3). These are of course "local times as space-indexed processes", up to normalization conventions. Occasionally I use "local time at a point" as a time-indexed process, still using the occupation density normalization.

Technical note. Using the construction of $\mathcal{S}$ from Brownian excursion, we get a random measure μ on $\mathcal{S}$ induced from Lebesgue measure on $[0,1]$: this is the same measure μ which occurs as the limit empirical distribution of vertices (section 2.3). The same applies to $(\mathcal{R}, \nu)$, the SSCRT: in the construction below from 2-sided BES(3), ν is the measure induced from Lebesgue measure on the line.

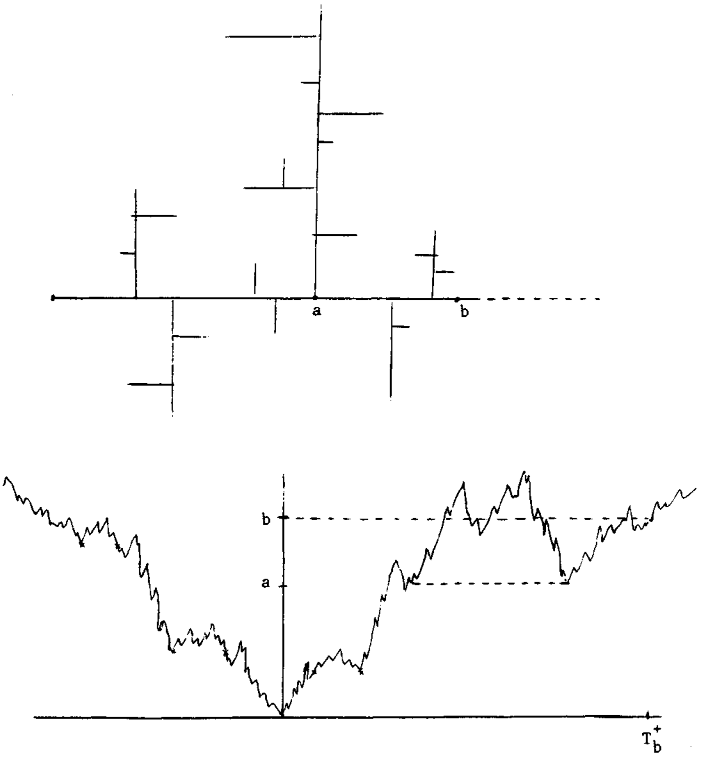

figure 2

2.5 The Self-Similar Continuum Random Tree

Sketched above is the SSCRT $(\mathcal{R}, \nu)$ given by the global construction in section 2.3. (The "baseline" is drawn horizontally.)

Recall that standard BES(3) is the process distributed as the radial part of 3-dimensional standard Brownian motion started at 0. We shall be concerned with 2-sided standard BES(3) $B = (B_s; -\infty < s < \infty)$. Here *2-sided* means that $(B_t; t \geq 0)$ and $(B_{-t}; t \geq 0)$ are independent copies of standard BES(3). It turns out that we can construct a realization of $\mathcal{R}$ from a realization of $2B$, analogous to the construction of $\mathcal{S}$ from $2W$ in section 2.2. In brief, we construct a tree labelled by $\{t : t \geq 0\}$ from $(2B_t; t \geq 0)$ and separately construct another tree labelled by $\{-t : t \geq 0\}$ from $(2B_{-t}; t \geq 0)$; then we join the trees by identifying (for each $b \geq 0$) the points labelled T_b^- and T_b^+, where

$$T_b^+ = \max\{t : 2B_t = b\},\ T_b^- = \min\{t < 0 : 2B_t = b\}.$$

This becomes the point $\mathbf{b}$ on the baseline at distance b from the root. In figure 2, we regard positive-time BES(3) as tracing out the part of the tree above the baseline, and negative-time BES(3) tracing out the part below the baseline.

Here is a verbal description of how $\mathcal{R}$ arises as a limit of rescaled CBP(n). Rescale edges to have length $1/a(n)$, where throughout this section

$$a(n) \to \infty,\ a(n) = o(n^{1/2}).$$

Let ν_n be the measure putting mass $1/a^2(n)$ on each vertex. Then the rescaled random set $\mathcal{T}_n$ of vertices of CBP(n) converges in distribution to $\mathcal{R}$, and the random measure ν_n converges to ν. From this limit procedure (or from the BES(3) construction) we see that the SSCRT has a self-similarity property: multiplying distances by a constant c doesn't affect the distribution of the random set $\mathcal{R}$, though it does take the measure ν to $c^{-2}\nu$.

These results could be formalized and proved in the same way as was done in [3] for the CCRT. Here are the concrete results analogous to Theorem 2 and Corollary 3.

Reconsider the search depth $(x(i); 1 \leq i \leq 2n-1)$ associated with a tree t in section 2.2. The search starts and ends at the root: $x(1) = x(2n-1) = 0$. For present purposes we want to center at the root, so we define $(x^*(i); -n < i < n)$ by

$$x^*(i) = x(i), 1 \leq i < n; \quad x^*(-i) = x(2n-i), 1 \leq i < n$$

with $x^*(0) = -1$. For a random tree distributed as CBP(n) this becomes a random sequence $(X^*(i))$; we also have the height profile process $(H(j))$ as in section 2.2. Rescale as

$$\hat{H}_s^n = a^{-2}(n) \sum_{j \leq a(n)s} H(j),\ s \geq 0 \tag{15}$$

and

$$\hat{X}_s^n = a^{-1}(n) X^*([2a^2(n)s]), -\infty < s < \infty. \tag{16}$$

Theorem 5 *For CBP(n), as n → ∞*

$$(\hat{X}_s^n; -\infty < s < \infty) \xrightarrow{d} (2\sigma^{-1} B_s; -\infty < s < \infty)$$

where B is 2-sided standard BES(3).

Corollary 6 *For CBP(n), as n → ∞*

$$(\hat{H}_s^n; s \geq 0) \xrightarrow{d} (Q_{\sigma s/2}; s \geq 0) \tag{17}$$

where

$$Q_s = \int_{-\infty}^{\infty} 1_{(B_t \leq s)} dt.$$

Here is the analog of Conjecture 4 for the (local) height profile process.

Conjecture 7 *For CBP(n), as n → ∞,*

$$(a^{-1}(n) H([a(n)s]);\ s \geq 0) \xrightarrow{d} (\frac{\sigma}{2} q_{\sigma s/2};\ s \geq 0) \overset{d}{=} (\frac{\sigma^2}{4} q_s;\ s \geq 0)$$

where

$$q_s = \frac{dQ_s}{ds}$$

is total occupation density for 2-sided BES(3).

Kolchin [33] Theorem 2.5.4 and Kennedy [30] Theorem 1 have given the 1-dimensional convergence results implicit in Conjecture 7, but I have not seen the full weak convergence result published explicitly.

It is well known that the total occupation density of one-sided BES(3) is the diffusion with drift and variance rates

$$\mu(x) = 2, \quad \sigma^2(x) = 4x,$$

or equivalently $|B_2(s)|^2$, where B_d is standard d-dimensional Brownian motion. It follows that (q_s) is $|B_4(s)|^2$, or equivalently the diffusion with

$$\mu(x) = 4, \quad \sigma^2(x) = 4x.$$

The marginal distributions are Gamma$(2, \cdot)$:

$$f_{q(s)}(x) = \frac{x}{4s^2}\exp(-\frac{x}{2s}). \tag{18}$$

See Pitman and Yor [40, 41] for extensive accounts of related properties of Bessel processes. As mentioned above, this Gamma limit distribution for generation size in conditioned Galton-Watson processes was known, but analytic proofs give little insight into why this particular limit distribution holds. The BES(3) representation gives one: the positive-time and negative-time occupation densities are obviously i.i.d. exponentials.

2.6 Discrete limits of CBP(n)

As a final piece of background, one can take limits in CBP(n) without rescaling edge-lengths. In this setting, the limit process $\mathcal{T}_\infty$ (described below) depends on the entire distribution of ξ. This result is in Grimmett [24] and in [2] Theorem 2, in the special case of Poisson(1) offspring; and the general case is implicit in Kesten [31].

The discrete infinite tree $\mathcal{T}_\infty$.
For each $k = 0, 1, 2, \ldots$ create independently branching processes, whose first generation size has distribution $\hat{\xi}$ but whose subsequent offspring distribution is ξ. Regard these as trees with root i_k and other vertices unlabelled. Then connect $i_0, i_1, i_2, \ldots$ as a path, deem i_0 the root and delete labels. □

This is a convenient place to introduce the idea of *random re-rooting.* A random tree $\mathcal{T}_n$ distributed as CBP(n) is normally considered as rooted at the progenitor of the branching process. We may, however, choose another vertex v of $\mathcal{T}_n$ and declare that to be the root. (To avoid discussing ordering of the re-rooted tree, regard trees as unordered). If v is chosen uniformly at random from the n vertices, call this procedure "random re-rooting". In the combinatorial model "uniform random labelled unordered tree", i.e. the Poisson(1) special case of CBP(n), it is immediate from the combinatorial description that random re-rooting does not change the distribution of the random tree. For general CBP(n), the distribution does change. However, the discrete limit distribution $\mathcal{T}^*_\infty$ is almost the same as $\mathcal{T}_\infty$ above, except for one change:

the branching process rooted at i_0 has first generation

offspring distribution ξ instead of $\hat{\xi}$. (19)

This result, implicit in earlier work, is given explicitly in Aldous [1]. As an aside, the idea of taking discrete limits in randomly re-rooted trees works for almost all the larger class of "height $O(\log n)$ trees" mentioned in the introduction, whereas for those trees looking at limits around the original root is not interesting – this topic is the subject of [1].

2.7 Symmetries of Trees, and the Arrow of Time

We now have four ways to look at the CCRT $\mathcal{S}$ (Brownian excursion, the global construction, limits of CBP(n), and the random f.d.d.'s (13)). An audience from theoretical stochastic processes is likely to concentrate on the first way, and think the whole subject is just a corner of Brownian excursion theory. But I hope to show that misses the point: all four ways are useful in doing calculations.

As an illustration, the fact that in a special case CBP(n) is exactly invariant under random re-rooting implies immediately that

the distribution of the CCRT is invariant under random re-rooting. (20)

By considering the search depth process, we could write this as a statement about Brownian excursion W. Fix u and define

$$\begin{aligned}\hat{W}_u(s) &= W_u + W_{u+s} - 2 \inf_{u \le t \le u+s} W_t, \ 0 \le s \le 1-u \\ &= W_u + W_{u+s-1} - 2 \inf_{u+s-1 \le t \le u} W_t, \ 1-u \le s \le 1.\end{aligned}$$

Then (20) becomes:

$$W \overset{d}{=} \hat{W}_U, \quad \text{where } U \text{ is uniform on } [0,1]. \qquad (21)$$

This is much less helpful than (20). I find it conceptually helpful to think of trees as purely **spatial** objects, without any notion of "time" involved. The point is that here are many different ways to associate "time" with a tree: the "intrinsic" time mentioned below if different from the notions of time induced by the Brownian excursion construction in section 2.2 and different again from the notions of time in the global constructions in section 2.3. Further, in section 5 we will consider trees as range spaces for random processes, in which setting having a notion of "time" attached to the tree itself is really confusing.

Having said this, recall that in a discrete-time branching process such as CBP(n) we would normally think of the vertices at distance d from the root as "the individuals alive at time d", since we are drawing the family tree with edges of unit length. Analogously, in a continuum tree we may consider "time" to be "distance from the root" – I call this *intrinsic time*. Loosely, we may think of the CCRT as a family tree for individuals with infinitesimal lifetimes, the vertices at distance t from the root representing the individuals alive at time t. Thus the processes (l_t) and (q_t) in the previous sections represent population sizes at time t in the CCRT and SSCRT. But the interesting symmetries of our trees, such as (20), involve changes in direction of intrinsic time, and this is why it helps to think of the trees as purely spatial objects.

As illustration, consider the *interpretation of the SSCRT as an ancestor process*. In section 2.5 the SSCRT was presented as a limit of rescaled CBP(n) as seen from the progenitor. Here the direction of time is indicated in the top diagram in figure 3. But as at (19) and (20) we can look at CBP(n) from the standpoint of a uniform random individual. Then rescaling as in section 2.5 gives the same limit SSCRT. Here the interpretation of the baseline is as the ancestral line back from the random individual V towards the progenitor, and a bush branching off the baseline at $\mathbf{b}$ indicates relatives of V whose last common ancestor with V was at (rescaled) time b in the past. See the middle diagram in figure 3. (Incidently, the bottom diagram arises in a context discussed in section 6.)

Relations between the limits. There are several relations amongst these processes.

1. $\mathcal{R}$ is the "large-scale" limit of $\mathcal{T}_\infty$ (the discrete infinite tree), and the "small-scale" local (i..e. around the root) limit of $\mathcal{S}$ ([2] Theorem 11). The latter fact is a translation of the fact that BES(3) is the rescaled limit of Brownian excursion near 0.

2. $\mathcal{R}$ can be obtained by attaching to the baseline a σ- finite process of (mostly

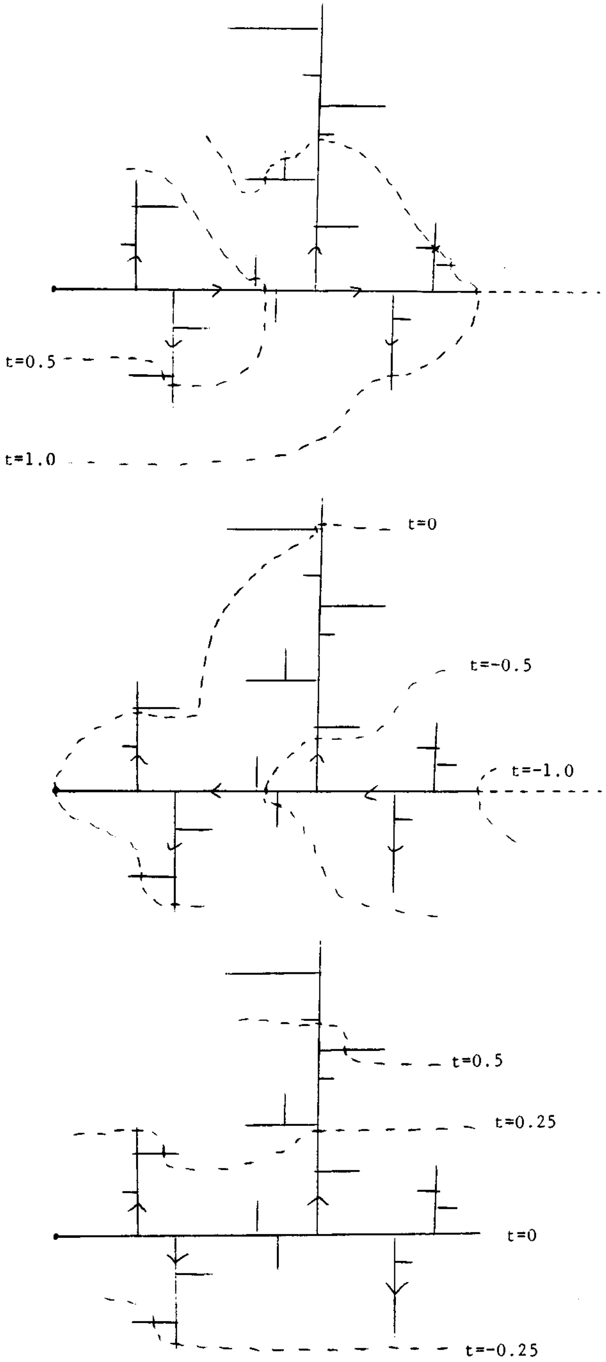

Figure 3

small) rescaled copies of $\mathcal{S}$ ([2] section 6). In figure 2, the "bush" attached at a arises from the excursion of $2B$ above a, drawn over the dashed line. This translates to a "last-exit" decomposition of BES(3) into excursions above levels b ending at the last exit time from b. See section 3.5 for applications.

3. The fact that Brownian excursion is "BES(3) bridge" is suggestive, but I see no solid interpretation in terms of our trees.

2.8 Discussion

The preceding sections contain my subjective view of "the big picture". But there is much more one could say about related matters.

1. In classical applied probability, there is a branching process description of the total number of customers served in a busy period of a $M/G/1$ queue. For a critical $M/M/1$ queue, this gives a correspondence between the continuous time simple symmetric random walk (number of arrivals – number of departures) and the shifted geometric (1/2) Galton-Watson branching process, and this is exactly the correspondence of section 2.2 translated into continuous time.

2. There is recent theoretical literature on trees associated with Brownian-type processes. Neveu and Pitman, whose work is summarized in [37], discuss trees associated with upcrossings of size h in Brownian excursion conditioned to reach height h (instead of conditioning on duration). The trees they obtain are the family trees of continuous-time critical branching processes where individual lifetimes have exponential($2/h$) distribution and are followed either by death (probability 1/2) or by a split into 2 new individuals (probability 1/2). Their construction resembles that in section 2.2 in that branchpoints correspond to local minima. But fundamental to my set-up is the idea of trees as having distances between vertices, and one really needs to draw the trees as in figure 1 to make this work.

3. Conversely, Waymire et al. ([25],[9] p. 284) start with the continuous-time binary branching process above and show that, conditioning on total population size $= n$ and letting $n \to \infty$, the time to extinction rescales to the maximum of Brownian excursion.

4. For the reader interested in pursuing distributional properties of Brownian excursion, relevant papers include Chung [12]; Knight [32]; Salminen [44]; Imhof [27]; Biane [10]; Biane and Yor [11].

5. As a fanciful analogy, there are two ways to paint a picture on a piece of paper. You can divide the paper into small pixels and paint each in turn; or you can start with broad brush strokes in the middle and then fill in medium and smaller size details. The latter is analogous to the global construction of the CCRT in section 2.3; the former is analogous to its construction from Brownian excursion, where the sample path of the excursion "traces the outline of the tree".

6. Obviously we could replace the CBP(n) model with the model of critical Galton-Watson branching processes conditioned to have *height* (i.e. number of generations before extinction) greater than h. Then a rescaled $h \to \infty$ limit is the variant of the CCRT obtained as in section 2.2 from Brownian excursion conditioned to reach height 1 at least. This seems less natural from the viewpoint of discrete random tree models. I do not know if this limit has a global construction like those of section 2.3, or a simple description of random subtrees like (13).

3 DISTRIBUTIONAL PROPERTIES

Obviously branching processes are very amenable to study via generating function methods. Various questions about CBP(n) have been studied by combinatorialists (and some probabilists) using exact formulae in special cases and generating function asymptotics for the general case. Kolchin [33] provides a useful summary of the extensive Russian work in this area. We shall see how well the "weak convergence, continuum trees and Brownian excursion" approach does on these questions.

3.1 Height

Write G_n for the *height* of CBP(n), i.e. the number of generations before extinction. Since G_n is the maximum of the search depth process, an obvious corollary of Theorem 2 is

Corollary 8 *For CBP(n), as $n \to \infty$*

$$n^{-1/2}\mathrm{G}_n \xrightarrow{d} 2\sigma^{-1}W^*$$

where $W^* = \sup_{0\leq t\leq 1} W_t$ is the maximum of Brownian excursion.

Expressions for the mean and the distribution of W^* are well known in the stochastic processes literature, e.g. Kennedy [29] or [9] p. 85:

$$EW^* = \sqrt{\pi/2} \tag{22}$$

$$P(W^* \leq x) = 1 - 2\sum_{k=1}^{\infty}(4x^2k^2 - 1)\exp(-2x^2k^2). \tag{23}$$

It is undoubtedly true that all moments converge in Corollary 8, but I did not keep track of moments in [3] so this does not rigorously follow from our approach. The result for means

$$n^{-1/2}E\mathrm{G}_n \to \sqrt{2\pi}\,\sigma^{-1} \tag{24}$$

was established via generating function asymptotics by Flajolet and Odlyzko [22], generalizing various special cases known earlier. The general limit distribution result of Corollary 8 is Theorem 2.4.3 of Kolchin [33]. Special cases have been known for a long time: Renyi and Szekeres [42] studied the "uniform random unordered labelled tree" and obtained an expression for the limit distribution which (using Corollary 8) becomes the expression

$$P(W^* \leq x) = 2^{1/2}\pi^{5/2}x^{-3}\sum_{k=1}^{\infty} k^2\exp(-k^2\pi^2/2x^2). \tag{25}$$

So the right sides of (25) and (23) must be equal. The special case of "uniform random ordered trees" (where of course the result is immediate from the ideas of section 2.2) has also been studied – see Takacs [49] for a recent treatment and references.

Remark. Here and elsewhere, combinatorial arguments typically give *local* limit theorems, which are stronger than the convergence in distribution obtained by our methods.

3.2 Height Profile

Corollary 3 and Conjecture 4 provide a connection between occupation density $(l_s; s \geq 0)$ of Brownian excursion W_t and asymptotic height profiles of CBP(n). In this section we look at the explicit formulas available in the literature.

Working directly with Brownian excursion, Knight [32] Theorem 2.3 gives the following expression for the marginal density of l_s.

$$f_{l_s}(y) = 2^{3/2}\pi^{5/2}s^{-3}\int_0^1 f_{W^*}\left(\frac{\pi^2(1-t)}{2s^2}\right)f(t,y)dt$$

where f_{W^*} is the density of W^* (i.e. the derivative of (23)), and where

$$f(t,y) = -\frac{(2\pi t)^{-1/2}}{2s}\sum_{i=0}^{\infty}\frac{1}{i!}\frac{d^{i-1}}{dy^{i-1}}\left(y^i\frac{d^2}{dy^2}\exp(-(2t^{-1}s^2(y+i)^2))\right)$$

Convergence of 1-dimensional distributions in the setting of Conjecture 4 follow from classical asymptotics for generation sizes and extinction times in critical branching processes: see Theorem 2.5.6 of Kolchin [33] or Kennedy [30] Theorem 3. This approach leads to the following indirect expression for the density.

$$f_{l_s}(y) = \frac{y}{4}\int_0^1 (1-t)^{-3/2} \exp\left(-\frac{y^2 s^2}{8(1-t)}\right) g_{2s}(y/2,t)\, dt$$

where $g_s(y,t)$ is the density whose joint characteristic function $\psi_s(\theta_1,\theta_2)$ is given by

$$1/\psi_s(\theta_1,\theta_2) = \frac{\sinh(s\sqrt{-2i\theta_2})}{s\sqrt{-2i\theta_2}} - i\theta_1\left(\frac{\sinh(s\sqrt{-i\theta_2/2})}{s\sqrt{-i\theta_2/2}}\right)^2.$$

Finally, by combinatorial analysis of the uniform random ordered tree, Takacs [47] obtains the formula

$$f_{l_s}(y) = 2\sum_{j=1}^{\infty}\sum_{k=1}^{j} \binom{j}{k} e^{-(y+2sj)^2/2} \frac{(-y)^k}{(k-1)!} H_{k+2}(y+2sj)$$

where H_k are the Hermite polynomials.

There are simpler formulas for moments, e.g. for means

$$El_s = 4s\exp(-2s^2)$$

but these are best though of as facts about the *distribution* of heights of *random* vertices of $\mathcal{S}$, as in section 3.3. Instead of emphasizing exact formulas (about which I have nothing new to say), let me emphasize some symmetry properties. In terms of CBP(n) with height G_n, there is no offspring distribution for which the height profile process exactly satisfies

$$(H(j); 0 \le j \le \mathrm{G}_n) \overset{d}{=} (H(\mathrm{G}_n - j); 0 \le j \le \mathrm{G}_n). \tag{26}$$

So from the branching process viewpoint there is no reason to suspect that the occupation time process (l_s) has the height-reversal symmetry property

$$(l_s; 0 \le s \le W^*) \overset{d}{=} (l_{W^*-s}; 0 \le s \le W^*). \tag{27}$$

But this is indeed true. Then Corollary 3 gives a sense in which (26) is always asymptotically true as $n \to \infty$. The symmetry (27) and a related identity

$$\sup_s l_s \overset{d}{=} 2W^*. \tag{28}$$

have some relevance to interesting questions about CBP(n). Here is one example: others are in the next section.

Odlyzko and Wilf [38] were interested in the maximal height profile

$$H_n^* = \max_j H(j)$$

for CBP(n). This is difficult to analyze by combinatorial methods, and required a lot of work to get a $O(n^{1/2}\log n)$ upper bound for EH_n^*. In view of (28), Conjecture 4 would imply

$$n^{-1/2}H_n^* \xrightarrow{d} \sigma W^*$$

and suggest the result for means

$$n^{-1/2}EH_n^* \to \sigma\sqrt{\pi/2}.$$

Finally, one could consider the sum $\sum_j jH(j)$ of heights of all n vertices of CBP(n). Corollary 3 implies

Corollary 9

$$n^{-3/2}\sum_j jH(j) \xrightarrow{d} 2\sigma^{-1}I$$

$$\textit{where } I = \int_0^1 W_s ds.$$

Darling [14] gives an expression for the Laplace transform of I. Takacs [48] gives a combinatorial proof of a special case of Corollary 9 and gives a complicated expression for the distribution of I in terms of infinite sums and special functions.

Heuristics for (27) and (28). These results are a small part of a big picture discussed in detail by Biane and Yor [11]. From my viewpoint they are anomalous because they do not seem to follow from any symmetry property of the continuum tree $\mathcal{S}$. Here are heuristics in terms of branching process asymptotics. Let U be the diffusion on state space $[0,\infty)$ with drift rate $\mu(x) = 0$ and variance rate $\sigma^2(x) = x$. This is the continuous limit of the generation size process in a (unconditioned) critical Galton-Watson branching process. More exactly, the limit where the initial population is $u_0 n^{1/2}$, the offspring variance is 1, the population size is divided by $n^{1/2}$ and the inter-generation time is $n^{-1/2}$. Thus the limit process in Conjecture 4 (with $\sigma = 1$) ought to be the conditioned diffusion

$$l^*_t \equiv (U_t; 0 \leq t < \infty | \int_0^T U_s \, ds = 1, U_0 = 0) \tag{29}$$

where $T = \inf\{t > 0 : U_t = 0\}$. Thus we are conditioning on U having an excursion from 0 of *area* 1.

With this description of l^* the "height-reversal symmetry" property (27) becomes intuitively obvious: a 1-dimensional diffusion is reversible, and conditioning on a reversible event preserves reversibility.

From Corollary 3 we have

$$(l^*_s; s \geq 0) \overset{d}{=} (\frac{1}{2} l_{s/2}; s \geq 0) \tag{30}$$

where l is the occupation density of Brownian excursion W. But there is another way of looking at l^*. Being a drift-free diffusion, U_t is a time-change of standard Brownian motion $\beta(t)$. With this time-change representation, a miracle occurs: the conditioning in (29) becomes conditioning β to have an excursion of *duration* 1, i.e. to be Brownian excursion W. Precisely, we get

$$(l^*_t; t \geq 0) \overset{d}{=} (W_{L^{-1}(t)}; t \geq 0) \text{ , where } L(u) = \int_0^u 1/W_s \, ds. \tag{31}$$

Putting together (31) and (30) gives a result of Jeulin (really a conditional form of the classic Ray-Knight description of local time for Brownian motion - see Bianne [10] Theorem 3) saying that the occupation density for Brownian excursion is a random time change of another Brownian excursion. And this relation gives (28).

3.3 Heights of Specified Vertices

Asking about asymptotics of heights $h_n(v)$ of particular vertices v in CBP(n) doesn't quite make sense: one has to specify how the vertex is chosen. Obviously Theorem 2 gives one case. Fix s and let v_n be the $[2ns]$'th vertex visited in the depth search process: then

$$n^{-1/2} h_n(v_n) \xrightarrow{d} 2\sigma^{-1} W_s.$$

The limit marginal density of Brownian excursion is given by the formula (Ito-McKean [28] section 2.9 (3a))

$$f_{W_s}(x) = 2^{1/2}\pi^{-1/2}s^{-3/2}(1-s)^{-3/2}x^2 \exp(-x^2/(2s(1-s))) \tag{32}$$

While the limit result (for general CBP(n)) is novel, this way of picking vertices is not particularly interesting from the viewpoint of discrete random trees. Instead, let us consider $h(V)$, where V is a random vertex of $\mathcal{S}$ chosen according to the "uniform" measure μ, and h denotes height. So $h(V) \overset{d}{=} 2W_U$, where U is uniform on $[0,1]$. As explained below, this has density

$$f_{h(V)}(x) = x\exp(-x^2/2) \tag{33}$$

and so

$$Eh(V) = \sqrt{\pi/2}. \tag{34}$$

So Theorem 2 implies the result for uniform vertices V of general CBP(n):

$$n^{-1/2}h_n(V) \xrightarrow{d} \sigma^{-1}h(V) \tag{35}$$

and suggests the result for all moments, in particular for means

$$n^{-1/2}Eh_n(V) \to \sigma^{-1}\sqrt{2/\pi}. \tag{36}$$

These limit results (35,36) for general CBP(n) were proved by generating function methods by Meir and Moon[35] Theorems 4.5 and 4.6 (in special cases, exact formulas are available). In fact they proved the local limit theorem corresponding to convergence of expectations of 1-dimensional distributions in Conjecture 4:

$$Eh_s^n \to \frac{\sigma}{2}El_{\sigma s/2} = \sigma^2 s\exp(-\sigma^2 s^2/2).$$

Note that implicit in (34) and (22) is the fact that the mean height of $\mathcal{S}$ is exactly twice the mean height of a random vertex of $\mathcal{S}$:

$$Eh(V^*) = 2Eh(V) = \sqrt{2\pi} \tag{37}$$

where V^* denotes a vertex at maximal height. An explanation of "exactly twice" comes from the stronger fact

$$E(h(V)|h(V^*)) = \frac{1}{2}h(V^*). \tag{38}$$

This follows from the height-reversal symmetry property (27) of the limit height profile process l^*, because $s \to l_s^*$ is the conditional density of $h(V)$ given $\mathcal{S}$.

There are several ways to understand (33), of which integrating (32) over $0 < s < 1$ is the least useful. The most elegant is to use the fact (14) that the subtree $\mathcal{R}(k)$ of $\mathcal{S}$ spanned by k uniform random vertices is distributed as the tree produced by the first k branches in the global construction of section 2.3. So $h(V_1)$ is distributed as the first cut-point in the global construction, which obviously has density (33). Properties of the joint distribution of $(h(V_1),\ldots,h(V_k))$ can in principle by obtained from the explicit distribution (13) of the subtree. For example when $k=2$, a tree with leaves at heights y_1, y_2 has edges of lengths x, y_1-x, y_2-x (for some $x \leq y_1 \wedge y_2$), and so we can use (13) to see that $(h(V_1), h(V_2))$ has joint density

$$f(y_1,y_2)=\int_0^{y_1\wedge y_2}(y_1+y_2-x)\exp(-(y_1+y_2-x)^2/2)\ dx.$$

Note that $f(s,s) = El_s^{*2}$, for the limit height profile l_s^* (i.e. for Brownian excursion occupation density, up to factors of two (30)). This provides some alternative explanations for formulas in Chung [12] section 6.

3.4 Diameter of the Compact Continuum Tree

The *diameter* Δ_n of CBP(n) is the maximal distance between a pair of vertices. The abstract result behind Theorem 2 (or Theorem 2 itself) implies

$$n^{-1/2}\Delta_n \xrightarrow{d} \sigma^{-1}\Delta \tag{39}$$

where Δ is the diameter of the CCRT $\mathcal{S}$. Using the representation of $\mathcal{S}$ in terms of Brownian excursion W,

$$\Delta = 2\sup_{0\leq t_1<t_2\leq 1}(W_{t_1}+W_{t_2}-2\inf_{t_1\leq t\leq t_2}W_t). \tag{40}$$

Szekeres [46] gave a generating function proof of the existence of a limit in (39) for the special case "uniform random unordered labelled trees". From his result we obtain the following expression for the density function $f_\Delta(x)$.

$$3/\sqrt{2\pi}\ f_\Delta(x) = \tag{41}$$

$$\sum_{m=1}^{\infty}\{\frac{64}{x^4}(4b_{m,x}^4-36b_{m,x}^3+75b_{m,x}^2-30b_{m,x})+\frac{8}{x^2}(4b_{m,x}^3-10b_{m,x}^2)\}\exp(-b_{m,x})$$

where $b_{m,x}=(8\pi m/x)^2$.

From this Szekeres computes

$$E\Delta=\frac{4}{3}\sqrt{2\pi}=\frac{4}{3}EG \tag{42}$$

where G is the height of the CCRT $\mathcal{S}$.

As facts about Brownian excursion, (41) and (42) are novel. It is an open problem to establish them directly from (40). Incidently, the argument for (41) is similar to the argument giving (25) for the limit height; so it is likely that (41) has an equivalent expression resembling (23) in format.

I want to present an informal "argument by symmetry" which explains the simple relation (42) between mean diameter and mean height. Given the continuum random tree $\mathcal{S}$, choose a point V uniformly in the tree (according to the measure μ) and let G^* be the height of the tree rooted at V. Then $\mathrm{G}^* \overset{d}{=} \mathrm{G}$ by re-rooting symmetry (20). I shall argue informally

$$E(\mathrm{G}^*|\Delta) = 3\Delta/4 \tag{43}$$

which obviously implies (42).

As a preliminary, although $\mu(\mathcal{S}) = 1$ by definition, we can think of "$\mathcal{S}$ conditioned on $\mu(\mathcal{S}) = c$" as the limit of CBP(cn) with $\sigma = 1$ under the $n^{-1/2}$ rescaling.

It is clear that $\mathcal{S}$ has a unique "center" v, that is a point such that $\mathcal{S}$ can be regarded as the union of two trees $\mathcal{S}_1, \mathcal{S}_2$ rooted at v and each having height $\Delta/2$. These trees have random sizes $(\mu(\mathcal{S}_1), \mu(\mathcal{S}_2)) = (A_1, A_2)$, say, where $A_1 + A_2 = 1$. The key fact is

$$\text{conditional on } (\Delta, A_1),\ \mathcal{S}_1 \text{ and } \mathcal{S}_2 \text{ are independent}$$
$$\text{and distributed as } \mathcal{S} \text{ conditioned on having}$$
$$\text{height} = \Delta/2 \text{ and size} = A_1 \text{ (resp. } A_2)$$

One sees this informally by considering the "uniform random unordered labelled tree" model with even diameter, where a "center" exists. In section 3.2 we discussed the height profile process l_s^* for $\mathcal{S}$ in terms of excursions of the diffusion U. Writing l^{1*} for the height profile process of $\mathcal{S}_1$ we get in the notation of (29)

$$\text{conditional on } (\Delta, A_1),\ (l_s^{1*}; 0 \le s \le \Delta/2) \overset{d}{=}$$
$$(U_s; 0 \le s \le \Delta/2)|T = \Delta/2, \int_0^{\Delta/2} l_s^{1*} ds = A_1).$$

But the conditioning preserves the time-reversibility of the excursions of U. Thus, conditional on (Δ, A_1),

$$A_1^{-1} E \int_0^{\Delta/2} s l_s^{1*} ds = \Delta/4$$

(c.f. the argument below (38)). This expression gives the conditional mean distance from the center to a point V_1 chosen uniformly in $\mathcal{S}_1$. Clearly the height of $\mathcal{S}$ rooted at V_1 is this distance plus $\Delta/2$. Applying the same result for $\mathcal{S}_2$ gives (43).

3.5 Processes Associated with the SSCRT

We now turn to the SSCRT $\mathcal{R}$ constructed globally in section 2.3, or from 2-sided BES(3) B_s in section 2.5. In section 2.5 we discussed the height profile process (q_s): here I shall discuss some distributions of other processes defined in terms of the tree $\mathcal{R}$. For $b > 0$ it is useful to write $\mathbf{b}$ for the point on the baseline at distance b from the root, and $\mathcal{R}_b$ for the part of $\mathcal{R}$ connected to the initial segment $[\text{root}, \mathbf{b}]$ of the baseline.

The projection process. This is the process $(Z(b); b \geq 0)$, where $Z(b) = \nu(\mathcal{R}_b)$, the total "weight" of $\mathcal{R}_b$. There are two interpretations of the process as limits in CBP(n), using either the original root or the random re-rooting procedure of section 2.7, and the latter is more interesting. Let V_n be a uniform random vertex of CBP(n). Let $V_n^*(b)$ be the ancestor of V in the $ba(n)$'th generation before V, and let $Z_n(b)$ be the total number of descendants of $V_n^*(b)$. Then from Theorem 5 (for re-rooted CBP(n))

$$\sigma^{-2}a^{-2}(n)Z_n(b) \xrightarrow{d} Z(b).$$

In [2] section 7 the global construction was used to prove

Lemma 10 *$(Z(b), b \geq 0)$ is the positive stable $(1/2)$ process, that is*

$$E\exp(-\theta Z(b)) = \exp(-b\sqrt{2\theta})$$

$$Z(b) \overset{d}{=} b^2 Z(1).$$

It is convenient to record an easy calculation here:

$$\int_0^b (b-s)Z(ds) \overset{d}{=} (4/9)b^3 Z(1). \tag{44}$$

One can alternatively obtain Lemma 10 from the BES(3) representation. The last exit time process $(T_b^+; b \geq 0)$ for $(2B_t)$ is a positive stable $(1/2)$ process (see e.g. [40]) and then $Z(b) \overset{d}{=} T_b^+ + T_b^-$ because ν is the measure induced by $2B$ from Lebesgue measure.

Remark. One could construct BES(3) by starting with the last exit time process (T_b^+) and then filling in excursions above levels b. In the global construction, each bush attached to the baseline represents such an excursion.

The depth process. Write F_b for the height of $\mathcal{R}_b$, considered rooted at the original root, and write D_b for the height of $\mathcal{R}_b$, considered re-rooted at $\mathbf{b}$. Think of D_b as the "depth" of $\mathbf{b}$. We can give D_b an interpretation as a limit in CBP(n), using as above a uniform random vertex V of CBP(n). Let $D_n(b)$ be the number of generations until extinction, for the process of descendants of the ancestor of V in the $ba(n)$'th generation before V. Then

$$a^{-1}(n)D_n(b) \xrightarrow{d} D_b.$$

A symmetry property baseline reversibility which is obvious from the global construction is the following: the distribution of $\mathcal{R}_b$ is invariant under reflection of the baseline segment [root, $\mathbf{b}$] about its midpoint. In particular

$$D_b \overset{d}{=} F_b \quad \text{for each } b. \tag{45}$$

But (F_b) and (D_b) are different as *processes*, e.g. because $D_b - b$ is non-decreasing in b whereas F_b does not have that property.

Lemma 11 *For each b,*

$$P(F_b \leq a) = P(D_b \leq a) = (1 - b/a)^2, \ a > b$$

This can be obtained from the description of F_b in terms of BES(3):

$$F_b = \sup_{T_b^- \leq s \leq T_b^+} 2B_s.$$

For by the hitting probability formula for BES(3)

$$P(\sup_{0 \leq s \leq T_b^+} 2B_s < a) = P(\inf_{s \geq 0} B_s > b/2 \mid B_0 = a/2) = 1 - b/a.$$

Last common ancestors. Any two vertices of $\mathcal{R}$ have a last common ancestor, the point at which the paths from the root to the vertices diverge. Questions like the following are natural in terms of the interpretation of the SSCRT as a limit of critical Galton-Watson branching processes conditioned to survive forever. Define C_b to be the distance from the root to the last common ancestor of *all* points at distance b from the root (this last common ancestor must be on the baseline). And define G_b to be the distance from the root to the last common ancestor of two randomly-chosen points V_1, V_2 at distance b from the root (chosen according to conditioned ν, the measure with total mass q_b) – this last common ancestor need not be on the baseline. These processes inherit the 1-self-similarity property

$$C_b \overset{d}{=} bC_1, \quad G_b \overset{d}{=} bG_1$$

as does F_b above.

Proposition 12 *(a)* $P(C_1 > c) = (1-c)^2, \ 0 < c < 1$.
(b) $P(G_1 > g) = 2g^{-2}(1-g)(g + (1-g)\log(1-g)), \ 0 < g < 1$.

To see (a), consider the point process $\mathcal{P} = \{(s^*, h^*)\}$ recording the heights h^* and positions s^* of bushes branching off the baseline in the global construction. Then $\mathcal{P}$ is a Poisson point process of intensity $\rho(s,h) = 2h^{-2}$. This fact comes out of the argument in [2] section 6, or alternatively from the BES(3) construction using excursions from last exit times. Plainly

$$\begin{aligned} P(C_1 > c) &= P(\text{ no points } (s,h) \text{ of } \mathcal{P} \text{ with } s < c \text{ and } h > 1-s) \\ &= \exp(-\int_0^c \int_{1-s}^{\infty} \rho(s,h)dhds) \end{aligned}$$

giving (a). For (b), we can use the 2-sided BES(3) description to rewrite G_1 as follows. Consider local time measure on $\{s : B_s = 1\}$; pick at random two times from this measure (normalized to a probability measure), and write $T_{(1)}, T_{(2)}$ for the order statistics of these two times; define

$$\begin{aligned} G_1 = &\ \min\nolimits_{t<T_{(1)} \text{ or } t>T_{(2)}} B_t \quad \text{if } T_{(1)} < 0 < T_{(2)} \\ = &\ \min\nolimits_{T_{(1)}<t<T_{(2)}} B_t \quad \text{if not.} \end{aligned}$$

Routine but tedious calculations with BES(3) lead to (b).

4 DIFFERENT MODELS FOR RANDOM TREES.

Notwithstanding the open questions mentioned in sections 2 and 3, I regard the story of the invariance principle for CBP(n) as now well-understood. From the viewpoint of asymptotics for discrete trees, there are two natural research directions.

Similar limits for more general models. I am willing to make a bold conjecture:

> In any reasonable model for random n-trees where the diameter is $\Theta(n^{1/2})$, the rescaled trees converge to limit processes which coincide with, or can be simply derived from, the limit trees discussed here.

Here is an example of a different model, IMST(n). Start with n isolated vertices. Repeatedly, choose a pair of vertices uniformly and join them by an edge, provided they are not already in the same component. Ultimately a tree is obtained.

The discrete infinite tree limit (c.f. section 2.6) for this model was obtained in Aldous [4], This limit is different from the CBP(n) limit, but rescaling the discrete limit tree into a continuum tree gives exactly the SSCRT. This is strong evidence – but nowhere near a proof - that $n^{-1/2}$ rescaling of IMST(n) gives the CCRT. The "exchangeability" formalizations in [3] are designed to help with examples like this. Loosely, all we need is an argument that

$$P(D_n > xn^{1/2}) \to \exp(-x^2/2) \tag{46}$$

where D_n is the distance in IMST(n) between two prespecified vertices, and then the techniques of [3] could bootstrap the argument into a proof of convergence to the CCRT. Unfortunately (46) seems difficult.

Here are some rather different models where we expect the CCRT limit.

1. Uniform random unordered unlabelled trees.
2. Uniform random spanning trees of expander graphs (e.g. hypercubes) - see Aldous [5].
3. Steele's [45] "exponential family" of random n-trees, with a parameter determining the mean proportion of leaves.

Another interesting application is to *random mappings*, a well-studied topic surveyed in Kolchin [33]. Here we choose uniformly at random one of the n^n functions $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$ and consider the graph with edges $(i, f(i))$. The component containing 1 consists of a cycle with attached trees: by representing the trees as in section 2.2, one can represent the entire mapping as a walk of length $2n$. In ongoing work with Jim Pitman it is shown that these walks rescale to reflecting Brownian bridge.

Different limit continuum random trees. A much harder topic is the study of asymptotics for random trees whose definition involves the geometry of d-dimensional space. Two combinatorial examples are the uniform random spanning trees of Z^d studied in Pemantle [39], and the Euclidean minimum spanning trees on random points in R^d studied in Aldous and Steele [7]. And numerous examples such as directed animals and DLA appear in the physics literature. In many of these examples it is natural to conjecture the existence of continuum limit trees after rescaling, but I am not aware of any rigorous proofs.

5 BROWNIAN MOTION ON CONTINUUM TREES.

5.1 Generalities

I want to discuss distributional properties of "standard Brownian motion" $(\mathbf{X}_t; t \geq 0)$ taking values in the CCRT $(\mathcal{S}, \mu)$ or in the SSCRT $(\mathcal{R}, \nu)$ (by default, random processes start at the root). The discussion is necessarily heuristic, because no rigorous proof of existence of $\mathbf{X}_t$ has been written down. Such processes may be interesting as counterparts to the recent rigorous theory of Brownian motion on regular fractals developed by Barlow and Perkins [8], Lindstrom [34] and others. Loosely, our particular CRTs have "dimension 2", inherited from Brownian sample paths: in view of the rigorous construction of continuum trees from general continuous functions in [3], one can certainly construct continuum trees of any fractional dimension. Despite the large physics literature, rigorous study of diffusions on non-regular fractal sets seems difficult. But obviously a tree structure is a great simplification, and a rigorous theory of Brownian motion on rather general continuum trees seems a natural next step. I outline below the shape that such a theory might take, but do not intend to pursue the topic myself.

It is important to regard $\mathcal{S}$ and $\mathcal{R}$ as *spatial* trees, and downplay their constructions from Brownian excursion and BES(3). For simple symmetric random walk on a discrete tree there is an elementary formula for the mean first passage time from one specified vertex to another. Using these formulas it is easy to see that in CBP(n) (with $\sigma^2 = 1$ for simplicity) the mean passage time between random vertices $\sim \sqrt{\pi/2}\, n^{3/2}$. Results of this type go back to Moon [36]. Thus one way to think of Brownian motion on $\mathcal{S}$ is as a rescaled limit of simple random walk on CBP(n): make the edges have length $n^{-1/2}$ so as to get the limit $\mathcal{S}$, and make the time between steps be $n^{-3/2}$. Similarly, we could start with simple symmetric random walk on the discrete infinite tree of section 2.6: the same rescalings should lead to Brownian motion on $\mathcal{R}$. The latter discrete setting was studied in Kesten [31], who refers to his unpublished work proving the existence of a limit $n^{-3/2}T_{n^{1/2}} \xrightarrow{d} T$ for the first passage time $T_{n^{1/2}}$ from the root to the point at distance $n^{1/2}$ along the infinite path. I interpret T as the time taken by $\mathbf{X}$ on $\mathcal{R}$ to first hit the point on the baseline at distance 1 from the root. The ingredients of a *proof* of the existence of **X** via such a weak convergence construction would plainly be similar to this unpublished work of Kesten.

In the next section I suggest an alternative "sample path" construction which I believe will handle more general continuum trees. But my main purpose is to exhibit concrete calculations of distributions (mostly expectations) associated with **X**, and this is the subject of sections 5.3 and 5.4.

5.2 An Occupation Density Construction

In this section let us work with the precise definition of a (deterministic) *continuum tree* $(\mathcal{S}_0, \mu_0)$ given in [3] section 2.3. Essentially, $\mathcal{S}_0$ is a set which is topologically a tree, i.e. has a unique non-self-intersecting path between any pair of points, and μ_0 is a probability measure on $\mathcal{S}_0$, which should be thought of as a "uniform measure". $\mathcal{S}_0$ contains a "root" denoted by 0. (As a technical remark, for our purposes here we also assume $\mathcal{S}_0 = \text{support}(\mu_0)$, though in [3] a weaker condition was used.)

We now define *Brownian motion* on a compact continuum tree $(\mathcal{S}_0, \mu_0)$ (the locally compact case needed for $\mathcal{R}$ is similar) to be a $\mathcal{S}_0$-valued process $(\mathbf{X}_t; t \geq 0)$ with the following properties.

(i) Continuous sample paths.

(ii) Strong Markov.

(iii) Reversible with respect to its invariant measure μ_0.

(iv) For each path $[[a,b]] \subset \mathcal{S}_0$ and each $x \in [[a,b]]$,

$$P_x(T_a < T_b) = \frac{d(x,b)}{d(a,b)} \tag{47}$$

where d denotes distance.

(v) For points $a, b \in \mathcal{S}_0$ let $m_{a,b}(\cdot)$ be the mean occupation measure for the process started at a and run until it first hits b. Then

$$m_{a,b}(dx) = 2d(b; c(b:a,x))\mu_0(dx), \;\; x \in \mathcal{S}_0 \tag{48}$$

where $c(b:a,x)$ is the point at which the paths $[[b,a]]$ and $[[b,x]]$ diverge.

The *skeleton* $\mathcal{S}_\infty$ of $\mathcal{S}_0$ is the set of points x which are in the interior of some path $[[a,b]]$. A consequence of the precise definition of continuum tree is that the skeleton has μ_0-measure zero. Thus the process spends Lebesgue-0 time in the skeleton.

In contrast to the setting of general fractal subsets of R^d [34], there is no difficulty in proving *uniqueness* here, because of the explicit formulas (47,48). A sledgehammer proof could be based upon the result about general Markov processes being determined up to random time-change by their exit place distributions.

To outline an existence argument, let $V_1, \ldots, V_k$ be picked independently from μ_0 and let $\mathcal{R}(k)$ be the subtree of $\mathcal{S}_0$ spanned by the root and the points $V_1, \ldots, V_k$. Then $\mathcal{R}(k)$ is a (random) tree consisting simply of a finite number of edges with positive edge-lengths. Recall (14) that in the particular case of the CCRT $\mathcal{S}$, this $\mathcal{R}(k)$ is distributed as the tree produced by the first k branches in the global construction of section 2.3. Let μ_k be the natural induced Lebesgue measure on $\mathcal{R}(k)$. Assume the following regularity property (*local homogeneity*): there exist deterministic $a_k \to \infty$ such that

$$\frac{1}{a_k}\mu_k \to \mu_0 \text{ a.s. as k } \to \infty \tag{49}$$

The CCRT has this property with $a_k = \sqrt{2k}$ (this is essentially Theorem 3 (ii) of [2], but also follows easily from the Brownian excursion representation).

On each $\mathcal{R}(k)$ we can define ordinary Brownian motion $X_k(t)$ started at the root 0. One way of viewing $\mathbf{X}$ is as the weak limit of these ordinary Brownian motions as we "fill out" the tree and speed up time:

$$X_k(a_k t) \xrightarrow{d} \mathbf{X}(t) \text{ as } k \to \infty.$$

In fact we can do better and use a sample path construction. The key idea is: if we measure time by a suitable "local time" rather than absolute time, then we can make these Brownian motions consistent as k increases. Fix $\tau > 0$. Run ordinary Brownian motion $X_k(t)$ on $\mathcal{R}(k)$ until the local time at the root reaches τ. Regard the accumulated local time as an occupation density l_k:

$$\int_0^\tau 1_{(X_k(t)\in A)}\, dt = \int_A l_k(\tau, \omega, x)\mu_k(dx), \ A \subset \mathcal{R}(k). \tag{50}$$

We can construct l_{k+1} in terms of l_k. For $\mathcal{R}(k+1)$ consists of $\mathcal{R}(k)$ plus a new edge (B_{k+1}, V_{k+1}) attached at a point $B_{k+1} \in \mathcal{R}(k)$. The occupation density $l_{k+1}(\tau, \omega, x)$ coincides with $l_k(\tau, \omega, x)$ for $x \in \mathcal{R}(k)$, while on the new edge its conditional distribution given $l_k(\tau, \omega, B_{k+1}(\omega)) = l$ is the occupation density for 1-dimensional Brownian motion on a line segment of length $d(B_{k+1}, V_{k+1})$

started at 0 and run until the local time at 0 reaches l. Continuing for all k, we get a function $l_\infty(\tau,\omega,x), x \in \mathcal{S}_\infty$ defined on the skeleton $\mathcal{S}_\infty$ of $\mathcal{S}_0$ which satisfies (50) for all k. If we can show

$$l_\infty(\tau,\omega,\cdot) \text{ extends to a bounded continuous function on } \mathcal{S}_0, \qquad (51)$$

then by (49) l_∞ is the $k \to \infty$ limit time-rescaled occupation density for Brownian motion on $\mathcal{R}(k)$, and this can be regarded as occupation density for some process $(\mathbf{X}(t) : 0 \le t \le t(\tau))$ on $\mathcal{S}_0$ run until its density at the root reaches τ, i.e. until absolute time

$$t(\tau) = \int_{\mathcal{S}_0} l_\infty(\tau,\omega,x)\mu_0(dx).$$

In other words, we are trying to construct $\mathbf{X}$ by specifying its occupation density up to times τ. The argument is rigorous up to (51); and it remains to put together different τ's and show that $\mathbf{X}$ has the properties (i)-(v). I conjecture that very little more than (49) is required – perhaps only some weak "metric entropy" condition on $\mathcal{S}_0$.

5.3 Easy Distributional Properties

We now return to the special cases of the CCRT $(\mathcal{S},\mu)$ and the SSCRT $(\mathcal{R},\nu)$, and assume that Brownian motion $\mathbf{X}$ exists as a weak limit of rescaled random walks as in section 5.1 and also via the construction in section 5.2.

The SSCRT case is somewhat more tractable, because it inherits from $\mathcal{R}$ a self-similarity property.

$$(\mathbf{X}_{ct}; t \ge 0) \stackrel{d}{=} (c^{1/3}\mathbf{X}_t; t \ge 0). \qquad (52)$$

One explanation of the "1/3" comes from the rescaling in section 5.1 of discrete random walk. Another comes from a scaling (53) of first hitting times, which we now derive. As in section 3.5, for $b > 0$ write $\mathbf{b}$ for the point on the baseline at distance b from the root, and write $Z(b) = \nu(\mathcal{R}_b) =$ the total "weight" of the part $\mathcal{R}_b$ of $\mathcal{R}$ connected to the initial segment $[\text{root}, \mathbf{b}]$ of the baseline. Appealing to (48)

$$E(T_{\mathbf{b}}|\mathcal{R}) = \int_0^b 2(b-y)Z(dy)$$

where we write $(\cdot|\mathcal{R})$ to mean conditioning on the realization of the random tree $(\mathcal{R},\nu)$. Now Lemma 10 says that $Z(\cdot)$ is the positive stable 1/2 process, and appealing to (44)

$$E(T_{\mathbf{b}}|\mathcal{R}) \stackrel{d}{=} \frac{8}{9}b^3 Z(1). \qquad (53)$$

Of course $EZ(1) = \infty$ and so the unconditional first hitting time T_b has infinite expectation: this is a disadvantage of working with the SSCRT.

Turning to the CCRT $(\mathcal{S}, \mu)$, consider first hitting times T_x on arbitrary $x \in \mathcal{S}$. Immediate from (48) is

$$E_x(T_y|\mathcal{S}) = \int_{\mathcal{S}} 2d(y, c(y : x, z))\mu(dz). \tag{54}$$

Another consequence of (48) is a simple formula for mean round trip times between leaves (here the root counts as a leaf)

$$E_x(T_y|\mathcal{S}) + E_y(T_x|\mathcal{S}) = 2d(x,y) \text{ for any leaves } x, y \in \mathcal{S}. \tag{55}$$

These have nothing to do with the particular structure of $\mathcal{S}$, but reflect elementary general identities for simple random walks on discrete trees. Now let V and V_1 be independent random points of $\mathcal{S}$ chosen according to μ (which puts mass 1 on the leaves – here continuum trees are simpler than discrete trees!). Using invariance (20) under random re-rooting, $Ed(V_1, V) = Ed(\text{root}, V)$ and $E_{V_1} T_V = ET_V (= E_{\text{root}} T_V)$. It follows from (55) that

$$\begin{aligned} ET_V &= Ed(\text{root}, V) \\ &= \sqrt{\pi/2} \text{ by (34)} \end{aligned} \tag{56}$$

which is the Brownian motion analog of the result of Moon [36] mentioned earlier. A slight elaboration is provided by

$$E(T_V|\ h(V)) = h(V) \tag{57}$$

where $h(V) = d(\text{root}, V)$. This is based upon another symmetry property of $\mathcal{S}$. Recall from (14) that we may regard $[[\text{root}, V]]$ as arising from the initial line-segment $[0, C_1]$ in the global construction of $\mathcal{S}$. Treating this initial line segment as a baseline, we can define a "projection process" $\hat{Z}(b)$ analogous to the SSCRT case (section 3.5). That is, $\hat{Z}(b)$ is the μ-mass ultimately attached to the first b units of the initial segment. The global construction makes clear the reversibility property

$$(\hat{Z}(b); 0 \le b \le h(V)) \stackrel{d}{=} (\hat{Z}(h(V) - b); 0 \le b \le h(V)). \tag{58}$$

Rewriting (54) as

$$E(T_V|\mathcal{S}) = \int_0^{h(V)} 2(h(V) - b)d\hat{Z}(b)$$

(58) implies (57).

It is immediate from (55) that, given $\mathcal{S}$, the v which maximizes the mean round trip time from the root to v and back is the vertex V^* at maximal distance from the root. For future reference,

$$E_{\text{root}} T_{V^*} + E_{V^*} T_{\text{root}} = 2\sqrt{2\pi}. \tag{59}$$

by (22). On the other hand, one can show that the v which maximizes the one-sided hitting time $E_{\text{root}}(T_v|\mathcal{S})$ is not V^*.

One could calculate variances of hitting times in similar ways, starting from formulas in the discrete setting. Moon's result in [36] Corollary 7.3.1 (whose proof relies on the special structure of the uniform random unordered tree) implies

$$\operatorname{var}(T_V) = \frac{32}{15}.$$

In principle this variance could be decomposed as the sum of three components – contributions from the choice of $\mathcal{S}$, the choice of V given $\mathcal{S}$, and the choice of Brownian path given $\mathcal{S}$ and V – but the computations look messy.

5.4 Hard Distributional Properties

The distribution of baseline hitting times for Brownian motion on the SSCRT turns out to be related to the distribution of inverse local time in the CCRT. Here is our best shot at describing these distributions, though it doesn't qualify as "explicit".

Consider 1-dimensional Brownian motion, started at 0 and run until its occupation density at 0 reaches a. It is well known ([43] VI.52) that its occupation density Z_s at position $s > 0$ behaves as the *Ray-Knight diffusion*

$$Z_0 = a, \ dZ_s = \sqrt{2Z_s} d\beta_s, \ s > 0, \tag{60}$$

where (β_s) is another 1-dimensional Brownian motion.

Now consider the global construction of $\mathcal{S}$ via the cut-and-join points (C_i, J_i). Fix $a > 0$ and define $(Z^a_s, s \geq 0)$ by

$$\begin{aligned} Z^a_0 &= a \\ dZ^a_s &= 2\sqrt{Z^a_s} d\beta_s \text{ on } C_{i-1} \leq s < C_i \\ Z^a_{C_i} &= Z^a_{J_i} \end{aligned}$$

Clearly $(Z^a_s, 0 \leq s \leq C_k)$ describes the occupation density over $\mathcal{R}(k)$ for Brownian motion on the tree $\mathcal{R}(k)$ constructed from the first k cuts, the motion run until the local time at the root 0 reaches a. It is not hard to argue (c.f. the urn model of [2] sec. 4) that the limits

$$L_a = \lim_{s_0 \to \infty} \frac{1}{s_0} \int_0^{s_0} Z_s^a \, ds \tag{61}$$

exist a.s. From the construction of Brownian motion $\mathbf{X}$ on $\mathcal{S}$ in section 5.2 we see that the process $(L_a, a \geq 0)$ is the "inverse local time at the root" process for $\mathbf{X}$. (This construction should be regarded as "unconditional on $\mathcal{S}$"). Of course, conditionally on $\mathcal{S}$ the process (L_a) is a subordinator, and $E(L_a|\mathcal{S}) = a$. Getting explicit distributional information about (L_a) seems the most important potentially solvable open question in this area.

Now consider the SSCRT. As in section 2.7 (relation 2) we may regard the SSCRT as a semi-infinite baseline with "bushes" attached, the bushes being rescaled copies of the CCRT. More precisely, let $(\mathcal{S}_c, \mu_c)$ denote the CCRT $(\mathcal{S}, \mu)$ after scaling by multiplying lengths by c and making the total measure $= c^2$. Then mark the baseline $[0, \infty)$ according to a marked Poisson process, marks in $[c, c + dc]$ appearing at rate $\sqrt{\frac{2}{\pi}} c^{-2} \, dc$. Wherever a mark c appears on the baseline, we attach a copy of $(\mathcal{S}_c, \mu_c)$.

Now consider Brownian motion $\mathbf{X}$ on the SSCRT, run until the time T_1 it first travels unit distance along the baseline. This has an occupation density (w.r.t. ν) process $(\hat{Z}(x), 0 \leq x \leq 1)$ on the baseline, which is just occupation density for 1-dimensional Brownian motion on the half-line started at 0 and run until first hitting 1. Specifically, $(\hat{Z}(1-x), 0 \leq x \leq 1)$ is the other Ray-Knight diffusion with drift rate $\mu(z) = z$ and variance rate $\sigma^2(z) = 2z$.

We are interested in the occupation density (over all $\mathcal{R}$) for $\mathbf{X}$ at time T_1. What happens on a bush depends only on the occupation density $\hat{Z}(x)$ at the point x where the bush attaches to the baseline. A scaling argument shows that inverse local time L_a^c for the Brownian motion on $(\mathcal{S}_c, \mu_c)$ scales as

$$(L_a^c, a \geq 0) \overset{d}{=} (c^3 L_{a/c}, a \geq 0).$$

Thus the amount of time spent in a bush with scaling factor c attached at x is distributed as $c^3 L_{\hat{Z}(x)/c}$. Adding over bushes gives

$$T_1 \overset{d}{=} \sum_{(x_i, c_i) \in \mathrm{POIS}} c_i^3 L^{(i)}_{\hat{Z}(x_i)/c_i} \tag{62}$$

where POIS is the Poisson process of points $(x, c) \in (0,1) \times (0, \infty)$ with rate $\sqrt{\frac{2}{\pi}} c^{-2}$ and where $L_a^{(i)}$ is distributed as at (61), independent as i varies. This is the advertized connection between SSCRT first passage times and CCRT inverse local time.

Changing topics, another quantity which (surprisingly) can be calculated explicitly is related to the *cover time*

$$C = \inf\{t : \mathcal{S} = \cup_{0\leq s\leq t}\mathbf{X}_s\}$$

for Brownian motion on $\mathcal{S}$, i.e. the first time at which the sample path has hit *every* point of $\mathcal{S}$. Consider the related cover-and-return time

$$C^+ = \inf\{t \geq C : \mathbf{X}_t = \text{root}\}.$$

Clearly C^+ is at least the time to visit the furthest leaf and return, which by (59) has mean $2\sqrt{2\pi}$. In fact I assert

$$EC^+ = 6\sqrt{2\pi}. \tag{63}$$

I do not know any simple symmetry argument for the factor of 3, though it is tempting to seek some overlooked symmetry. Pedestrianly, one can study random walk on unconditioned Galton-Watson trees and set up a recursion for the analogous cover-and-return time. Write C_n^+ for this time, conditioned on the size of the tree $= n$. In the case of Poisson offspring (i.e. the uniform unordered random labelled tree) it is proved in [6] that

$$\text{if } EC_n^+ \sim cn^{3/2} \text{ then } c = 6\sqrt{2\pi}. \tag{64}$$

This is an intuitively convincing argument for (63). At the rigorous level, a major issue is that, even granted weak convergence of rescaled random walks on discrete trees to the limit Brownian motion on $\mathcal{S}$, this does *not* imply convergence of cover times.

6 SUPERPROCESSES

This section is directed at readers already familiar with the subject of superprocesses. I have no technical knowledge of the subject, but am merely aiming to set out in conversational style some remarks about what is intuitively obvious, given a "random tree" background.

There is an underlying nice continuous-time Markov process $(X_t; t \geq 0)$ taking values in a space Σ. The associated superprocess takes values in the set $M(\Sigma)$ of non-negative measures on Σ, starting at time 0 with (say) the unit mass at a point x_0. It can be constructed as a weak limit of finite-population branching Markov processes, or directly via martingale characterization. Recently attention has been given to the associated "historical process" which gives the family tree of the limit population: see e.g. Dawson and Perkins [15], Dynkin [18, 19], Le Gall [23]. My viewpoint is to take this as the starting place: think about tree-indexed Σ-valued processes rather than time-indexed $M(\Sigma)$-valued processes.

Given a rooted tree t with a finite number of edges of positive length, let $\mathbf{t}$ be the point-set of all points of the tree, i.e. the points in the edges as well as the branchpoints and endpoints. Given a starting position x_0 for the underlying Markov process, there is an obvious construction of a tree-indexed Markov process $(X_s; s \in \mathbf{t})$, as follows. Put $X_{\text{root}} = x_0$ and then define X on one edge at a time, working away from the root. For an edge $[[a,b]]$ for which $X_a = x_a$ has already been defined, we define $(X_s, s \in [[a,b]])$ to be distributed as the underlying process started at x_a and run for time $d(a,b)$, independently of the previously-defined parts of the tree-indexed process.

Now consider a general continuum tree $(\mathcal{S}_0, \mu_0)$ (e.g. constructed from some function f as in section 2.2). By the Kolmogorov extension theorem we can define a tree-indexed process $(\mathbf{X}_s; s \in \mathcal{S}_0)$ such that for each finite set of points $v_1, \ldots, v_k$ in $\mathcal{S}_0$ the process restricted to the subtree $\mathbf{t}$ spanned by (v_i) is distributed as specified above. Of course such constructions by extension are unsatisfactory because different "versions" may have different sample path properties. To get a cleaner construction, suppose the underlying process has cadlag paths. Then we can construct $(\mathbf{X}_s; s \in \text{skeleton}(\mathcal{S}_0))$ such that every realization is cadlag at each point in the skeleton. For leaves $x \in \mathcal{S}_0$ one can seek to define by continuity:

$$\mathbf{X}_x = \lim\{\mathbf{X}_s : s \in [[0,x]],\ s \to x\}$$

In general the limit will exist outside a μ_0-null set of leaves, on which we must let $\mathbf{X}$ be undefined.

With each $s \in \mathcal{S}_0$ we associate the "intrinsic time" $t(s)$ which is just the distance from the root to s:

$$t(s) = d(\text{root}, s).$$

We can then define a time-indexed $M(\Sigma)$-valued process

$$\Theta_t(\cdot) = \int_{\mathcal{S}_0} 1_{(\mathbf{X}_s \in \cdot)} 1_{(t(s) \le t)} \mu_0(ds). \tag{65}$$

If sufficiently smooth in t we can differentiate to get

$$\theta_t(\cdot) = \frac{d}{dt}\Theta_t(\cdot). \tag{66}$$

Some notation: for a positive measure θ, write $||\theta||$ for its total mass.

Now consider this construction of a tree-indexed process $\mathbf{X}$ in the two special cases where the indexing tree is the CCRT $(\mathcal{S}, \mu)$ or the SSCRT $(\mathcal{R}, \nu)$. I claim these are the same as the usual superprocess, but with different conditionings. In the second case (the immortal superprocess) we are conditioning on the process being created (with infinitesimal mass) at point x_0 at time 0, and on the process surviving for ever. In the first case (the superprocess excursion) we are conditioning on the process being created (with infinitesimal mass) at point x_0 at time 0, and on the total (i.e. integrated over time up to extinction) population size being equal to 1. Arguably these conditionings are more natural in many biological applications than the usual "start with unit mass" model. To fit this usual model into our set-up, we use as index the continuum random tree pictured on the bottom in figure 3, i.e. the part of the SSCRT which branches off the first unit of the baseline, measuring "time" as "distance from baseline". In all these examples, what is usually called the superprocess is the $M(\Sigma)$-valued process (θ_t) at (66).

The assertions above are intuitively obvious from the interpretation of $\mathcal{S}$ and $\mathcal{R}$ as limits of family trees in critical branching processes. As with the development of historical processes, the point is to "decouple" the family tree structure from the Markov motion in the space Σ. In the next section I make some observations about superprocesses based on this "continuum random tree" viewpoint.

6.1 Five Observations

1. *Computer simulation.* Suppose you want to estimate some distribution associated with some explicit superprocess by computer simulation. The naive way would be to simulate a discrete-time critical branching process. Starting with 50, say, individuals and running until extinction would require order $50^2 = 2,500$ calls to the random number generator just to simulate the "family tree". It is much more efficient to use the global constructions in section 2.3 and simulate the first 50, say, branches of the CCRT, which requires only $2 \times 50 = 100$ calls to the generator (and then finally simulate the Markov process along the edges).

2. *Different models for random trees.* Consider the IMST(n) model of section 4. This is a random tree on n vertices, where we pay no attention to the order in which edges were added in the construction. Now imagine one vertex placed at $0 \in R^d$ and the edges having length $n^{-1/2}$ and independent uniform random directions in R^d. Let Φ^k be the empirical distribution of

the vertex positions. Granted the conjecture that this model also rescales to $\mathcal{S}$, it is intuitively obvious that Φ^k converges in distribution to Θ_∞, the total occupation density (65) associated with the superprocess excursion built over d-dimensional Brownian motion. The point is that the discrete IMST(n) model has no notion of "Markovian branching" or even of "time", but still leads to a superprocess.

3. *A connection between superprocesses and Brownian motion on continuum trees.* One quantity of interest in the latter context was L_a at (61), the time at which Brownian motion on the CCRT has accumulated local time a at the root. It is intuitively clear that

$$L_a \overset{d}{=} \int_{\mathcal{S}} \mathbf{X}_s^{(a)} \, \mu(ds)$$

where $\mathbf{X}^{(a)}$ is the $\mathcal{S}$-indexed process (i.e. superprocess excursion) built over the Ray-Knight diffusion (60).

4. *Symmetry properties.* The symmetry properties of $\mathcal{S}$ and $\mathcal{R}$ we have discussed lead to symmetry properties of superprocesses. Here is one example. Suppose the underlying Markov process is ergodic with stationary distribution π. Choose X_0 from π, then run the superprocess excursion starting at X_0. Let Θ_∞ be the total occupation measure (65) and pick X^* according to Θ_∞. If the underlying Markov process is reversible then

$$(X_0, \Theta_\infty, X^*) \overset{d}{=} (X^*, \Theta_\infty, X_0).$$

In words, the distribution of Θ_∞ relative to a random individual in the population is the same as its distribution relative to the progenitor: so the progenitor is not special. This property is intuitively clear from the "random re-rooting" property (20) of the CCRT.

5. *The immortal superprocess.* This has recently been studied rigorously by Evans and Perkins [21, 20]. From our viewpoint of the superprocess as the Markov process $\mathbf{X}$ indexed by the SSCRT, some elementary properties are obvious.

(a) The "population size at time t" $|\theta_t|$ has Gamma$(2, 2/t)$ distribution, by Corollary 6 and (18).

(b) If the underlying Markov process has absorbing states then (because the SSCRT consists of bounded bushes attached to an infinite baseline) the superprocess gets absorbed with the same probabilities as for the underlying Markov process ([20]).

(c) If the underlying Markov process converges from any start to a unique stationary distribution π then

$$\theta_t/|\theta_t| \xrightarrow{p} \pi. \tag{67}$$

For using a standard exchangeability fact (that for exchangeable sequences, pairwise independence implies independence) to prove (67) it suffices to prove

$$\text{dist}(\mathbf{X}_{V_1(t)}, \mathbf{X}_{V_2(t)}) \to \pi \times \pi \tag{68}$$

where $V_1(t)$ and $V_2(t)$ are picked uniformly from the population at time t. The individuals $V_1(t)$ and $V_2(t)$ had last common ancestor at time G_t, say, and

$$P(\mathbf{X}_{V_1(t)} \in A, \mathbf{X}_{V_2(t)} \in B | G_t = g, \mathbf{X}_{G_t} = x) = P_x(\mathbf{X}_{t-g} \in A) P_x(\mathbf{X}_{t-g} \in B).$$

By the self-similarity property of the SSCRT we have $G_t \overset{d}{=} tG_1$, and then (68) follows easily from the convergence assumption on the underlying Markov chain [21].

In fact, the exact distribution of G_1 was calculated in Proposition 12, as was the distribution C_1 of the time of the last common ancestor of the entire time-1 population.

References

[1] D.J. Aldous. Asymptotic fringe distributions for general families of random trees. 1991. To appear in Annals of Applied Probability, May 1991.

[2] D.J. Aldous. The continuum random tree I. *Ann. Probab.*, 19:1–28, 1991.

[3] D.J. Aldous. The continuum random tree III. In preparation.

[4] D.J. Aldous. A random tree model associated with random graphs. *Random Structures and Algorithms*, 1:383–402, 1990.

[5] D.J. Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.*, 3:450–465, 1990.

[6] D.J. Aldous. Random walk covering of some special trees. *J. Math. Analysis Appl.*, 157:xxx, 1991. To appear.

[7] D.J. Aldous and J.M. Steele. Asymptotics for Euclidean minimal spanning trees on random points. 1991. Unpublished.

[8] M.T. Barlow and E.A. Perkins. Brownian motion on the Sierpinski gasket. *Probab. Th. Rel. Fields*, 79:543–624, 1988.

[9] R.N. Bhattacharaya and E.C. Waymire. *Stochastic Processes with Applications.* Wiley, 1990.

[10] P. Biane. Relations entre pont et excursion du mouvement Brownien reel. *Ann. Inst. Henri Poincare*, 22:1–7, 1986.

[11] P. Biane and M. Yor. Valeurs principales associees aux temps locaux Browniens. *Bull. Sci. Math. (2)*, 111:23–101, 1987.

[12] K.L. Chung. Excursions in Brownian motion. *Arkiv fur Matematik*, 14:155–177, 1976.

[13] E. Csaki and G. Mohanty. Excursion and meander in random walk. *Canad. J. Statist.*, 9:57–70, 1981.

[14] D.A. Darling. On the supremum of a certain Gaussian process. *Ann. Probab.*, 11:803–806, 1983.

[15] D.A. Dawson and E. Perkins. Historical processes. 1991. Mem. Amer. Math. Soc., to appear.

[16] R. Durrett, D.L. Iglehart, and D.R. Miller. Weak convergence to Brownian meander and Brownian excursion. *Ann. Probab.*, 5:117–129, 1977.

[17] R. Durrett, H. Kesten, and E. Waymire. On weighted heights of random trees. *J. Theoretical Probab.*, 4:223–237, 1991.

[18] E.B. Dynkin. *Branching Particle Systems and Superprocesses.* Technical Report, Cornell University, 1989.

[19] E.B. Dynkin. *Path Processes and Historical Processes.* Technical Report, Cornell University, 1989.

[20] S.N. Evans. *The Entrance Space of a Measure-Valued Markov Branching Process Conditioned on Non-extinction.* Technical Report, U.C.Berkeley, 1991. Canad. Math. Bull., to appear.

[21] S.N. Evans and E. Perkins. Measure-valued Markov branching processes conditioned on non-extinction. *Israel J. Math.*, 71:329–337, 1990.

[22] P. Flajolet and A. Odlyzko. The average height of binary trees and other simple trees. *J. Comput. System Sci.*, 25:171–213, 1982.

[23] J.-F. Le Gall. *Brownian Excursions, Trees and Measure-Valued Branching Processes.* Technical Report, Universite P. et M. Curie, Paris, 1989.

[24] G. R. Grimmett. Random labelled trees and their branching networks. *J. Austral. Math. Soc. (Ser. A)*, 30:229–237, 1980.

[25] V.J. Gupta, O.J. Mesa, and E. Waymire. Tree dependent extreme values: the exponential case. *J. Appl. Probab.*, 27:124–133, 1990.

[26] W. Gutjahr and G. Ch. Pflug. *The Asymptotic Contour Process of a Binary Tree is Brownian Excursion.* Technical Report, University of Vienna, 1991. To appear in Stochastic Proc. Appl.

[27] J.-P. Imhof. On Brownian bridge and excursion. *Studia Sci. Math. Hungar.*, 20:1–10, 1985.

[28] K. Ito and H.P. McKean. *Diffusion Processes and their Sample Paths.* Springer, 1965.

[29] D.P. Kennedy. The distribution of the maximum Brownian excursion. *J. Appl. Prob.*, 13:371–376, 1976.

[30] D.P. Kennedy. The Galton-Watson process conditioned on the total progeny. *J. Appl. Prob.*, 12:800–806, 1975.

[31] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincare Sect. B*, 22:425–487, 1987.

[32] F. B. Knight. On the excursion process of Brownian motion. *Trans. Amer. Math. Soc.*, 258:77–86, 1980.

[33] V.F. Kolchin. *Random Mappings.* Optimization Software, New York, 1986. (Translation of Russian original).

[34] T. Lindstrom. Brownian motion on nested fractals. *Memoirs of the A.M.S.*, 420, 1989.

[35] A. Meir and J.W. Moon. On the altitude of nodes in random trees. *Canad. J. Math.*, 30:997–1015, 1978.

[36] J.W. Moon. Random walks on random trees. *J. Austral. Math. Soc.*, 15:42–53, 1973.

[37] J. Neveu and J. Pitman. The branching process in a Brownian excursion. In *Seminaire de Probabilites XXIII*, pages 248–257, Springer, 1989. Lecture Notes in Math. 1372.

[38] A.M. Odlyzko and H.S. Wilf. Bandwidths and profiles of trees. *J. Combin. Theory Ser. B*, 42:348–370, 1987.

[39] R. Pemantle. *Choosing a Spanning Tree for the Integer Lattice Uniformly.* Technical Report, Cornell University, 1989. To appear in Ann. Probability.

[40] J.W. Pitman and M. Yor. Bessel processes and infinitely divisible laws. In *Stochastic Integrals*, pages 285–370, Springer, 1981. Lecture Notes in Math. 851.

[41] J.W. Pitman and M. Yor. A decomposition of Bessel bridges. *Z. Wahrsch. Verw. Gebiete*, 59:425–457, 1982.

[42] A. Renyi and G. Szekeres. On the height of trees. *J. Austral. Math. Soc.*, 7:497–507, 1967.

[43] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales.* Wiley, 1987.

[44] P. Salminen. Brownian excursions revisited. In *Seminar on Stochastic Processes 1983*, pages 161–187, Birkhauser Boston, 1984.

[45] J.M. Steele. Gibb's measures on combinatorial objects and the central limit theorem for an exponential family of random trees. *Prob. Engineering Inf. Sci.*, 1:47–60, 1987.

[46] G. Szekeres. Distribution of labelled trees by diameter. In *Combinatorial Mathematics X*, pages 392–397, Springer-Verlag, 1982. Lecture Notes in Math. 1036.

[47] L. Takacs. Limit theorems for random trees. 1991. Unpublished.

[48] L. Takacs. On the total height of a random planted trivalent plane tree. 1991. Unpublished.

[49] L. Takacs. Queues, random graphs and branching processes. *J. Appl. Math. and Simulation*, 1:223–243, 1988.